The Philosophical Magazine

UNIVERSITY OF HAWAII
LIBRARY
Dec 13 '61

FIRST PUBLISHED IN 1798

A Journal of Theoretical Experimental and Applied Physics

Vol. 6 August 1961 No. 68
Eighth Series

25s. 0d., plus postage
Annual Subscription £13 10s. 0d., payable in advance

ALERE FLAMMAM

Printed and Published by
TAYLOR & FRANCIS LTD
RED LION COURT, FLEET STREET, LONDON, E.C.4

THE PHILOSOPHICAL MAGAZINE

Editor

Professor N. F. MOTT, M.A., D.Sc., F.R.S.

Editorial Board

Sir LAWRENCE BRAGG, O.B.E., M.C., M.A., D.Sc., F.R.S.
Sir GEORGE THOMSON, M.A., D.Sc., F.R.S.
Professor A. M. TYNDALL, C.B.E., D.Sc., F.R.S.

AUTHORS wishing to submit papers for publication in the Journal should send manuscripts directly to the Publishers.

Manuscripts should be typed in *double* spacing on one side of quarto (8 × 10 in.) paper, and authors are urged to aim at absolute clarity of meaning and an attractive presentation of their texts.

References should be listed at the end in alphabetical order of authors and should be cited in the text in terms of author's name and date. Diagrams should normally be in Indian ink on white card, with lettering in soft pencil, the captions being typed on a separate sheet.

A leaflet giving detailed instructions to authors on the preparation of papers s available on request from the Publishers.

Authors are entitled to receive 25 offprints of a paper in the Journal free of charge, and additional offprints can be obtained from the Publishers.

The *Philosophical Magazine* and its companion journal, *Advances in Physics*, will accept papers for publication in experimental and theoretical physics. The *Philosophical Magazine* publishes contributions describing new results, letters to the editor and book reviews. *Advances in Physics* publishes articles surveying the present state of knowledge in any branch of the science in which recent progress has been made. The editors welcome contributions from overseas as well as from the United Kingdom, and papers may be published in English, French and German.

Dislocations in Visco-elastic Materials

By J. D. Eshelby
Department of Physical Metallurgy, University of Birmingham

[Received March 23, 1961]

Abstract

An expression is found for the force required to move a dislocation steadily through a visco-elastic medium, in particular Zener's standard linear solid. It is applied to the case of a screw dislocation in iron exhibiting the Snoek effect. Schoeck and Seeger's atomic treatment of the same problem contains an error. When this is corrected the two methods give solutions of the same form.

---

§ 1. Introduction

Zener (1948) has given the name 'standard linear solid' to a simple type of visco-elastic material in which stress σ, strain e and their rates of change are connected by the equation

$$\sigma+\tau_1\dot{\sigma}=M_{\mathrm{R}}(e+\tau_2\dot{e}). \qquad (1)$$

If unit strain is suddenly imposed the stress is $M_{\mathrm{U}}=\tau_2 M_{\mathrm{R}}/\tau_1$ to begin with (unrelaxed modulus). It then falls to a lower value M_{R} (relaxed modulus) with a time of relaxation τ_1. If unit stress is suddenly applied the strain rises from $1/M_{\mathrm{U}}$ to $1/M_{\mathrm{R}}$ with a relaxation time $\tau_2>\tau_1$. The quantity

$$\Delta=\frac{M_{\mathrm{U}}-M_{\mathrm{R}}}{M_{\mathrm{R}}}=\frac{\tau_2-\tau_1}{\tau_1}$$

measures the 'strength' of the relaxation process. If σ and e both depend on time through a factor exp $(i\omega t)$ the relation (1) is equivalent to

$$\sigma=\mathscr{M}(\omega)e \qquad (2)$$

where

$$\mathscr{M}(\omega)=M_{\mathrm{R}}\frac{1+i\omega\tau_2}{1+i\omega\tau_1} \qquad (3)$$

is a complex elastic modulus. If $\mathscr{M}(\omega)$ has a more general form (subject to the requirement that it always leads to positive strain-energy and positive rate of dissipation) (2) describes the behaviour of a general linear visco-elastic solid (Hunter 1960).

It may seem rather academic to consider the movement of dislocations in such materials. However, a number of solid-state phenomena can be described in terms of relaxed and unrelaxed elastic moduli and a relaxation time. The Snoek effect (re-orientation of carbon atoms in alpha-iron) is the classic example, and there are others of the same kind. Certain effects involving phonon or electron relaxation can be fitted into the same picture. In § 2 we find an expression for the force required to keep a dislocation moving uniformly through the visco-elastic medium. Section 3 treats the particular case of a screw dislocation moving through iron showing the Snoek effect. We find it necessary to treat the non-linear region near the dislocation in a rather unsatisfactory way. Schoeck and Seeger's (1959) atomic treatment of the same problem in principle handles this region satisfactorily, but it contains an error. When this is corrected (§ 4) the result agrees in form with the visco-elastic solution, though the numerical value is different, and probably more reliable.

§ 2. RETARDING FORCE ON A DISLOCATION MOVING IN A VISCO-ELASTIC MEDIUM

Consider a linear visco-elastic solid in which the displacement is everywhere parallel to the z-axis and of the form $w=W(x, y)\exp(i\omega t)$. The only non-vanishing stress components are

$$\sigma_{zx}=\mathscr{M}(\omega)\partial w/\partial x, \quad \sigma_{zy}=\mathscr{M}(\omega)\partial w/\partial y$$

where $\mathscr{M}(\omega)$ is the complex shear modulus. If the motion is slow enough for the inertial forces to be neglected (we assume this throughout) the equilibrium equation is

$$\mathscr{M}(\omega)\nabla^2 w=0. \qquad (3\,a)$$

Let us impress on the face $y=0$ of the semi-infinite block $y\geqslant 0$ a sinusoidal corrugation $w=\exp\{-i(kx-\omega t)\}$ moving along the x-axis with constant velocity $v=\omega/k$. Experience with this kind of problem suggests that the displacement will penetrate into the block according to the law

$$w=\exp(-ly-ikx') \qquad (4)$$

where

$$x'=x-vt, \quad v=\omega/k.$$

Equation (3 a) is in fact satisfied if we set l equal to k. On $y=0$ the velocity of the material is $\dot{w}=i\omega\exp(-ikx')$ and all the stress components vanish except for σ_{zy} which has the value $\sigma=-k\mathscr{M}(\omega)\exp(-ikx')$. Cancelling a factor $i\omega$ and taking the real part we find that to produce a velocity-distribution

$$\dot{w}=\cos kx' \qquad (5)$$

on the surface of the block we must apply a stress

$$\sigma=v^{-1}\operatorname{re}\{\mathscr{M}(kv)\exp(ikx')\} \qquad (6)$$

to it. If we build up a more general surface velocity in the form of the Fourier integral

$$\dot{w}=\int_0^\infty f(k)\cos kx'\,dk \qquad (7)$$

the stress necessary to maintain it is evidently

$$\sigma=v^{-1}\int_0^\infty f(k)\{\operatorname{re}\mathscr{M}(kv)\sin kx'-\operatorname{im}\mathscr{M}(kv)\cos kx'\}\,dk. \qquad (8)$$

(We note in passing that by (4) $\dot{w}(x', y)$ is given by (7) with a factor $\exp(-k|y|)$ inserted in the integrand. Since this is independent of $\mathscr{M}(\omega)$, $w(x', y)$ is the same as it would be for a disturbance moving in a non-relaxing medium.) The work done by the surface forces in unit time is

$$\dot{E} = -2\int_0^\infty \sigma(x')\dot{w}(x')\,dx'. \tag{9}$$

Now take the block $y<0$ and impress a velocity distribution equal and opposite to (7) on its surface. The stresses match those on the surface of the block $y \geqslant 0$ and the two may be stuck together. If w is chosen suitably we shall have a moving screw dislocation. The total rate of dissipation of energy in the two blocks is $2\dot{E}$ and the retarding force on the dislocation is $F=2\dot{E}/v$. According to Parseval's theorem (Titchmarsh 1937) if

$$g_m(x') = \int_0^\infty G_m(k)\cos kx'\,dk, \quad m=1, 2$$

then

$$\int_0^\infty g_1(x')g_2(x')\,dx' = \tfrac{1}{2}\pi\int_0^\infty G_1(k)G_2(k)\,dk.$$

Thus from (7), (8) and (9)

$$F(v) = -\frac{2\pi}{v^2}\int_0^\infty f^2(k)\,\mathrm{im}\,\mathscr{M}(kv)\,dk. \tag{10}$$

Equation (10) applies to a general visco-elastic medium. If $\mathscr{M}(\omega)$ has the simple form (3) appropriate to a standard linear solid with $M_\mathrm{R}=\mu_\mathrm{R}$ it becomes

$$F(v) = \frac{2\pi}{v}\mu_\mathrm{R}(\tau_2-\tau_1)\int_0^\infty \frac{kf^2(k)}{1+k^2v^2\tau_1{}^2}\,dk.$$

If we take the usual expression

$$w = \frac{b}{2\pi}\tan^{-1}\frac{x'}{\zeta}$$

for the displacement at the slip plane of a dislocation of width ζ then

$$\dot{w} = -\frac{vb}{2\pi}\frac{\zeta}{x'^2+\zeta^2} = -\frac{vb}{2\pi}\int_0^\infty \exp(-k\zeta)\cos kx'\,dk$$

and

$$F(v) = \frac{\mu_\mathrm{R}(\tau_2-\tau_1)b^2}{2\pi v\tau_1{}^2}\int_0^\infty \frac{\exp(-2k\zeta)k\,dk}{k^2+(v\tau_1)^{-2}}$$

$$= \frac{\mu_\mathrm{R}b^2}{4\pi}\frac{\Delta}{\zeta}\mathscr{F}\left(\frac{v\tau_1}{2\zeta}\right) \tag{11}$$

where (see figure)

$$\mathscr{F}(z) = z^{-1}\{-\cos(z^{-1})\mathrm{Ci}(z^{-1}) - \sin(z^{-1})\mathrm{si}(z^{-1})\} \tag{12}$$

in the notation of Jahnke and Emde (1945). We have

$$\begin{aligned}\mathscr{F}(z)&=z, && z\ll 1\\ &=z^{-1}\ln(z/\gamma), && z\gg 1 \qquad (\ln\gamma=0{\cdot}5772\ldots),\\ \mathscr{F}_{\max}&=0{\cdot}35 \text{ at } z=1{\cdot}3,\end{aligned}$$

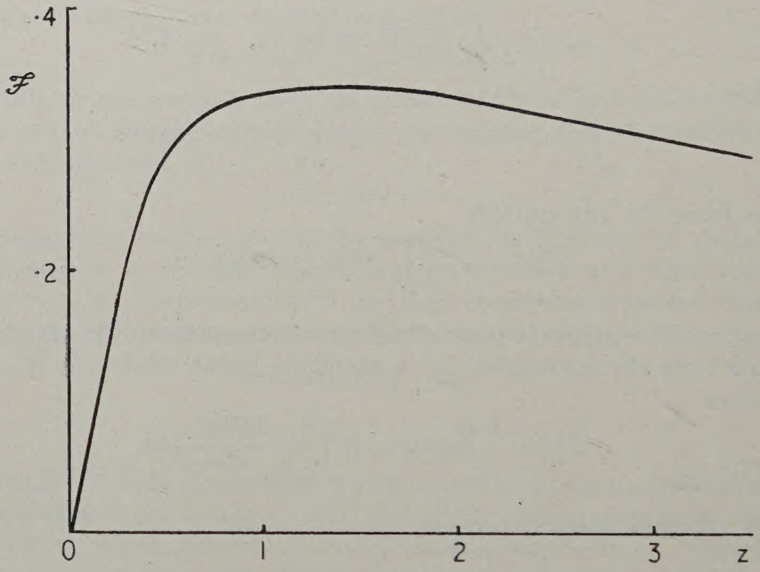
and so for low velocities the force is proportional to v,

$$F=\frac{\mu_{\mathrm{R}}}{8\pi}\frac{b^2}{\zeta^2}\tau_1\Delta v, \quad v\ll 2\zeta/\tau_1 \qquad . \quad . \quad . \quad (13)$$

while beyond a maximum

$$F_{\max}=\frac{0{\cdot}35}{4\pi}\frac{\mu_{\mathrm{R}}b^2\Delta}{\zeta} \quad \text{at } v_{\max}=1{\cdot}3(2\zeta/\tau_1)$$

it falls off with increasing v.



Following Mason (1960) we may derive the low velocity limit (13) directly. Far from the moving dislocation the Fourier components of its field will have frequencies small enough for us to write

$$\mathscr{M}(\omega)=\mu_{\mathrm{R}}[1+i\omega(\tau_2-\tau_1)]$$

or equivalently

$$\sigma=\mu_{\mathrm{R}}e+\eta\dot{e} \qquad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

where

$$\eta=\mu_{\mathrm{R}}\tau_1\Delta$$

is a small viscosity in parallel with the stiffness μ_{R}. The rate of flow of energy into unit volume is $\sigma\dot{e}=\mu_{\mathrm{R}}e\dot{e}+\eta\dot{e}^2$. The second term represents a dissipation at the rate $\eta\,(\mathrm{grad}\,\dot{w})^2$ per unit volume, whereas the first term corresponds to stored energy which is recovered if e goes through a cycle and returns to its initial value. According to the remark following (8) we may take $w=(b/2\pi)\tan^{-1}(x'/y)$ far from the dislocation. The total rate of dissipation outside a circle of radius r_0 is thus

$$\dot{E}=\frac{\eta v^2}{4\pi}\frac{b^2}{r_0^{\,2}} \quad . \; . \; . \; . \; . \; . \; . \; . \quad (15)$$

if (14) is a good approximation for $r>r_0$. Suppose that physical considerations (not directly connected with the visco-elastic behaviour of the medium) dictate an inner cut-off radius r_0. Then for small enough velocities (14) will be a good approximation for all $r\geqslant r_0$ and (15) is the whole dissipation. The retarding force is equal to (13) with $r_0=2\zeta$.

The expression (11) is related to the reduction in the energy of a stationary dislocation due to relaxation. If the dislocation is suddenly introduced its energy is $(\mu_{\mathrm{U}}b^2/4\pi)\ln(r_1/r_2)$ to begin with and $(\mu_{\mathrm{R}}b^2/4\pi)\ln(r_1/r_2)$ when relaxation is complete. The energy loss is

$$E_{\mathrm{rel}}=\tfrac{1}{2}W_0\ln(r_1/r_2) \quad . \; . \; . \; . \; . \; . \; . \quad (16)$$

where

$$W_0=(\mu_{\mathrm{U}}-\mu_{\mathrm{R}})b^2/4\pi$$

and we can write

$$F(v)=\frac{W_0}{2\zeta}\mathscr{F}\left(\frac{v\tau_1}{2\zeta}\right). \quad . \; . \; . \; . \; . \; . \; . \; . \quad (17)$$

Equations (16) and (17) can be generalized to cover the case of any straight dislocation (screw, edge or mixed) in an anisotropic medium with relaxed and unrelaxed elastic constants c_{ij}^{R}, c_{ij}^{U}, subject to two limitations. These are (i) all the c_{ij} have the same relaxation time τ_1, though the individual relaxation strengths $\Delta_{ij}=(c_{ij}^{\mathrm{U}}-c_{ij}^{\mathrm{R}})/c_{ij}^{\mathrm{R}}$ may be different, and (ii) the width parameter ζ has the same value for the edge and screw components of the dislocation. The elastic energy of the dislocation can be written in the form

$$E=\frac{Kb^2}{4\pi}\ln\frac{r_1}{r_2}$$

where K is a function of the direction of the Burgers vector and of the c_{ij}. In some cases K can be found immediately from the c_{ij}, in others it may be necessary to solve a cubic equation (Foreman 1955). The energy relaxation and force are given by (16) and (17) with

$$W_0=(K_{\mathrm{U}}-K_{\mathrm{R}})b^2/2\pi$$

where K_{U}, K_{R} are the values of K calculated from the c_{ij}^{U} and the c_{ij}^{R} respectively. The proof consists in writing out the relations analogous to (7) and (8) which connect displacement and stress on the face of a semi-infinite block, calculating the rate of dissipation as before and comparing the result with Foreman's expression for the energy as an integral over the slip-plane.

If both relaxation strengths and relaxation times are different for the various elastic constants things are more complicated. Consider, for example, an edge dislocation in an isotropic medium. The relation between displacement and stress at the surface of a semi-infinite block is now of the form (7, 8) with

$$\mathscr{M}(\omega)=\frac{\mu(\omega)}{1-\nu(\omega)}=2\mu(\omega)\frac{3\kappa(\omega)+\mu(\omega)}{3\kappa(\omega)+4\mu(\omega)} \quad . \quad . \quad . \quad (18)$$

where $\mu(\omega)$, $\kappa(\omega)$, $\nu(\omega)$ are the complex shear modulus, bulk modulus and Poisson's ratio. This is just a generalization to the visco-elastic case of a result given by Nabarro (1947). So long as we are concerned with a general visco-elastic material the expression (10) for the force is in principle no more complicated than it would be for a screw dislocation with $\mathscr{M}(\omega)=\mu(\omega)$. We might, however, be interested in a material which behaves as a standard linear solid under either shear or compression, but with a different relaxation time and strength for each of these types of deformation. Then $\mathscr{M}(\omega)$ (eqn. (18)) no longer has the simple form (3) and $F(v)$ depends on v through a more complicated function than the $\mathscr{F}$ of eqn. (11). It can easily be evaluated. If Δ and τ_1 happen to be the same for shear and compression, Poisson's ratio is independent of frequency and $F(v)$ is equal to (11) divided by $1-\nu$.

§ 3. Snoek Drag on a Screw Dislocation

Carbon atoms in a crystal of alpha-iron can occupy positions at the mid-points of 100, 010 or 001 cube edges. If a stress is applied in, say, the 100 direction there is a certain initial extension. Subsequently the crystal extends further as carbon atoms jump from 010 and 001 positions into the 100 positions. If the stress had been applied in the 111 direction there would have been no relaxation. Thus the material behaves as an anisotropic standard linear solid (Snoek effect). Polder (1945) gives the value

$$\tau_1=4{\cdot}4\times10^{-14}\exp(9000/T)\ \text{sec}$$

for the relaxation time. This is about one second at room temperature. A carbon atom need only jump one interatomic spacing to change from, say, an 010 to a 100 position, and no long-range diffusion is involved. Indeed it is often convenient to ignore the change of position altogether and think of the carbon atom as a ‘ dipole ’ which can take up one of three orientations. By analogy with paramagnetism, Kröner (1958) has given the name ‘ paraelastic ’ to materials which behave as visco-elastic solids because of the re-orientation of atomic dipoles of some kind. Zener (1948) has suggested that in some circumstances the interaction between the carbon atoms may be enough to make them all take up the same orientation spontaneously and so form martensite. If this is in fact so, iron containing carbon may reasonably be called a ferroelastic material.

Nabarro (1948) pointed out that (apart from forming carbide particles on them) carbon atoms in iron can interact with dislocations in two ways. They may re-orient to form a 'Snoek atmosphere'. This is a rapid process. Given time they may migrate over large distances and form a Cottrell atmosphere. The difference between the effects of the Snoek and Cottrell atmospheres is nicely demonstrated by Wilson and Russell's (1959) experiments on strain-ageing in low carbon steels.

We shall apply the results of § 2 to find the retarding effect of the Snoek atmosphere about a uniformly moving screw dislocation.

According to Polder (1945), the relaxed and unrelaxed moduli for alpha-iron containing an atomic fraction c of carbon are related by

$$c_{11}{}^{\mathrm{R}}=c_{11}{}^{\mathrm{U}}-2a,\ c_{12}{}^{\mathrm{R}}=c_{12}{}^{\mathrm{U}}+a,\ c_{44}{}^{\mathrm{R}}=c_{44}{}^{\mathrm{U}} \quad . \ . \ (19)$$

with

$$a=\tfrac{1}{6}(c_{11}{}^{\mathrm{U}}-c_{12}{}^{\mathrm{U}})\beta^2a_0{}^3c/kT$$

where a_0 is the lattice parameter, $\beta=1{\cdot}00$ is the rate of change of the axial ratio of martensite with carbon content and k, T are Boltzmann's constant and the absolute temperature. (As we ignore interaction between the carbon atoms we put Polder's constant χ equal to zero.)

Following Schoeck and Seeger (1959) we consider a screw dislocation in the 111 direction. This is not one of the cases for which K_{R} and K_{U} can be found easily. Since in any case we shall not be able to determine the length ζ very accurately we shall simply average the relaxed and unrelaxed anisotropic moduli and treat the problem as an isotropic one. If, with Voigt, we average the c_{ij} rather than the s_{ij} we obtain

$$\mu_{\mathrm{U}}-\mu_{\mathrm{R}}=\tfrac{3}{5}a.$$

With Polder's figures this gives

$$\Delta=1{\cdot}9\times10^4\,c/T. \quad . \ . \ . \ . \ . \ . \ . \ (20)$$

Zener (1948) gives $\Delta=2{\cdot}5\times10^4\,c/T$ for extension in the 100 direction, for which the effect is strongest.

We have next to choose the length ζ. For a given ζ the maximum shear strain due to the dislocation is $b/4\pi\zeta$. In an ordinary elastic problem we should choose ζ so that the maximum strain does not exceed the value for which Hooke's law becomes invalid. This would be wrong in the present calculation. If we apply a slowly increasing stress to the material the shear modulus is initially μ_{R}. Above a certain stress, however, nearly all the carbon atoms will have entered the favoured sites, and thereafter the rate of increase of stress with strain is nearly equal to μ_{U}, and anelastic effects are greatly reduced. Thus we must choose ζ in such a way that the maximum strain $b/4\pi\zeta$ does not exceed the value at which, so to speak, Hooke's law breaks down for the difference-modulus $\mu_{\mathrm{R}}\Delta=\mu_{\mathrm{U}}-\mu_{\mathrm{R}}$. At normal temperatures this is much less than the strain at which Hooke's law breaks down for μ_{U} or μ_{R}. Zener (1948) gives a curve (p. 122) which shows how the number of carbon atoms in preferred positions varies with strain and temperature. The curve becomes markedly non-linear at a strain of about 0·005 at room temperature, or a strain 0·005 $T/300$ at absolute temperature T. We shall therefore require ζ to satisfy

$$b/4\pi\zeta = 2\times10^{-5}\,T. \qquad \qquad (21)$$

From (11), (20) and (21) we find that the stress required to maintain a velocity v is

$$\sigma = F(v)/b = 3\times10^{11}\,c\mathscr{F}(v\tau_1/2\zeta)\ \text{dynes/cm}^2 \qquad (22)$$

or very nearly $10^{11}\,c$ dynes/cm² at $v_{\max}$.

§ 4. THE METHOD OF SCHOECK AND SEEGER

In the previous section our treatment of the anelastically non-linear region near the dislocation was very crude. Schoeck and Seeger (1959) have already discussed the same problem. Their treatment, which is atomistic, runs into no special difficulties near the dislocation. However, it contains an error. To show this we first give a brief account of their method. The error is easily corrected, and the result is then substantially the same as ours.

Schoeck and Seeger begin by calculating the interaction energy between a stationary dislocation and the equilibrium Snoek atmosphere surrounding it. The result is

$$U_0 = W_0 \ln(3L/R)$$

where

$$W_0 = \frac{2\pi c}{a_0^3}\frac{A^2}{kT},$$

$$R = \frac{A}{kT}$$

and

$$A = 1{\cdot}84\times10^{-20}\ \text{dyne-cm}^2.$$

The length L is a cut-off radius which is taken to be 10^{-4} cm. The factor 3 in the argument of the logarithm stands for exp (1·1). U_0 is approximately twice our expression (16) because it is an interaction energy, not a self-energy. They then develop an ingenious method designed to avoid having to calculate the precise form of the Snoek atmosphere round the moving dislocation. For this purpose they define a function $V(t)$ which is " the energy a fictitious dislocation at the origin would acquire in the field of the impurity atoms ordered by a dislocation line at x " and another function $V_0(x)$ which is the energy the fictitious dislocation would acquire if the atmosphere about the real dislocation at x were in complete equilibrium (as it would be, for example, if the real dislocation had been at rest at x for a long time). They show that $V(t)$ and $V_0(x)$ are related by the equation

$$\frac{\partial V(t)}{\partial t} = -\bar{\nu}\,[V(t) - V_0(x)]$$

where $\bar{\nu}$ is an atomic jump frequency. For the force on a uniformly moving dislocation they find

$$F(v) = \left[\frac{1}{v}\frac{\partial V}{\partial t}\right]_{t=0} = \frac{\bar{\nu}}{v}\left[V_0(0) - \int_0^\infty V_0(vt)\exp(-\bar{\nu}t)\,dt\right]. \qquad (23)$$

Since the actual form of $V_0(x)$ is not known Schoeck and Seeger assume that it will be adequate to insert in (23) a trial function which satisfies two conditions: (i) $V_0(0) = U_0$, (ii) $V_0(x)$ falls rapidly to zero outside the range $|x| \sim R$. They actually take

$$V_0(x) = U_0 \exp(-x^2/R^2). \qquad . \quad . \quad . \quad . \quad . \quad . \quad (24)$$

This gives a force which can be written in the form

$$F(v) = \frac{W_0}{R}\ln\left(\frac{3L}{R}\right)\mathscr{G}\left(\frac{2v}{\bar{\nu}R}\right) \qquad . \quad . \quad . \quad . \quad . \quad . \quad (25)$$

where $\mathscr{G}(z)$ is a function similar to our $\mathscr{F}(z)$ with the following properties:

$$\mathscr{G}(z) = z, \ z \ll 1; \ \mathscr{G}(z) = 2z^{-1}, \ \ z \gg 1;$$

$$\mathscr{G}_{\max} = \tfrac{1}{2} \ \text{at} \ z = \tfrac{4}{3}.$$

If we suppose that $R \sim \zeta$ and $\bar{\nu} \sim \tau_1^{-1}$ (25) gives roughly the same velocity dependence as (11). Numerically (25) is about six times greater than our expression (11). There is nothing unreasonable in this. However, the disagreement is more than numerical. Equation (25) contains a factor $\ln(3L/R)$, estimated by Schoeck and Seeger to be $41/2\pi = 6{\cdot}5$. Our expression contains no such logarithmic cut-off factor. Indeed the discussion in § 2 leading to (15) shows that far from the dislocation the rate of dissipation per unit volume falls off as r^{-4}, and so it is difficult to see why an outer cut-off should be necessary.

In fact the logarithmic factor only appears because (24) (or any other function of short range) is a physically inadmissible trial function. Under moderate stresses the material behaves as an anisotropic standard linear solid with relaxed and unrelaxed moduli $c_{ij}{}^{\mathrm{R}}$, $c_{ij}{}^{\mathrm{U}}$ related by (19). Thus at large distances from the dislocation the difference between its elastic fields in the relaxed and unrelaxed condition must be the same as the field of a dislocation in an ordinary elastic medium with constants $c_{ij}{}^{\mathrm{U}} - c_{ij}{}^{\mathrm{R}}$. The interaction energy $V_0(x)$ of this difference-field with a fictitious dislocation at a distance x from it must thus behave like $\ln x$ for large x. Only a detailed calculation could give the value of $V_0(x)$ for small x, but by definition we must have $V_0(0)=U_0$. A trial function which meets these requirements is

$$V_0(x)=\tfrac{1}{2}W_0 \ln \frac{L^2}{x^2+r_0^2} \qquad . \ . \ . \ . \ . \ . \quad (26)$$

with

$$r_0=\tfrac{1}{3}R=A/3kT.$$

If this is inserted in (23) we get

$$F(v)=\frac{W_0}{r_0}\mathscr{F}(v/\bar{\nu}r_0)$$

or

$$\sigma=F/b=6{\cdot}0\times 10^{11}c\mathscr{F}(v/\bar{\nu}r_0). \qquad . \ . \ . \ . \ . \ . \quad (27)$$

With $\bar{\nu}=\tau_1^{-1}$ and $r_0=2\zeta$ the stress is twice as large as our (22), but the variation with velocity is the same. (Actually r_0 is about four times the value of 2ζ given by (21).)

Schoeck and Seeger compare the result of their calculation with a value

$$\sigma=4500c \text{ kg/mm}^2$$

which they deduce from experiment. (c is the atomic fraction of carbon.) Their original expression (25) gives $\sigma=F_{\max}/b=6300c$ kg/mm^2, but the revised expression (27) only gives $2000c$ kg/mm^2. Our estimate (22) gives $1000c$ kg/mm^2. We can get better agreement if we identify the experimental value not with the force needed to attain the velocity $v_{\max}$, but rather with the force, F^* say, required to pull the dislocation away from its Snoek atmosphere suddenly, so that there is no time for the carbon atoms to move. Schoeck and Seeger find that $F^*=F_{\max}$. But it is easy to see that F^* is simply the maximum value of $\partial V_0(x)/\partial x$. If $V_0(x)$ has the form (26) then

$$F^*=W_0/r_0=2{\cdot}9F_{\max}$$

which gives $\sigma=6000c$ kg/mm^2. A rough calculation for the force to pull an edge dislocation suddenly from its Snoek atmosphere gives $\sigma=2500c$ kg/mm^2 (Eshelby 1959: the quantity ΔY must be divided by 2 to convert to a shear stress. On p. 631, col. 2, l. 20, $\gamma=\frac{1}{2}$ should read $\gamma=\frac{1}{3}$).

These values cannot be taken very seriously, since they are sensitive to the detailed form of $V_0(x)$ near the origin. For example, it is easy to show from (23) that

$$F(v)=-\bar{\nu}^{-1}V_0''(0)v+O(v^2)$$

for small v. Obviously a knowledge of the value of $V_0(x)$ at $x=0$ and of its behaviour at infinity tells us nothing about its curvature at the origin. $F_{\max}$ and F^* are somewhat less sensitive, but it is clear that really reliable values could only be obtained by calculating $V_0(x)$ in the same way as Schoeck and Seeger calculated $V_0(0)$.

References

ESHELBY, J. D., 1959, Appendix to WILSON, D. V., and RUSSELL, B., *Acta Met.*, **7**, 628.

FOREMAN, A. J. E., 1955, *Acta Met.*, **3**, 322.

HUNTER, S. C., 1960, *Progress in Solid Mechanics*, **1**, 1.

JAHNKE, E., and EMDE, F., 1945, *Tables of Functions* (New York: Dover Publications).

KRÖNER, E., 1958, *Kontinuumstheorie der Versetzungen und Eigenspannungen.* (Berlin: Springer–Verlag).

MASON, W. P., 1960, *J. acoust. Soc. Amer.*, **32**, 458.

NABARRO, F. R. N., 1947, *Proc. phys. Soc., Lond.*, **59**, 256; 1948, *Report on Strength of Solids* (London: Physical Society), p. 38.

POLDER, D., 1945, *Philips Res. Rep.*, **1**, 5.

SCHOECK, G., and SEEGER, A., 1959, *Acta Met.*, **7**, 469.

TITCHMARSH, B. C., 1937, *Fourier Integrals* (Oxford: Clarendon Press).

WILSON, D. V., and RUSSELL, B., 1959, *Acta Met.*, **7**, 628.

ZENER, C., 1948, *Elasticity and Anelasticity of Metals* (Chicago: University Press).

Hydrogen Nuclei of the Primary Cosmic Radiation

By C. J. Waddington
H. H. Wills Physical Laboratory, University of Bristol

[Received April 21, 1961]

Abstract

A small stack of nuclear emulsions has been used to determine the flux of cosmic ray protons over Texas. The experimental procedure was similar to that used previously, but due to the more favourable conditions of the exposure the dependence of the value at the top of the atmosphere on the, rather uncertain, corrections for the overlying matter, was greatly reduced, the correction being 0·902 c.f. the previous 0·694. The value found for the flux at the top of the atmosphere on July 12, 1960 was 540 ± 50 protons/m^2 ster sec.

§ 1. Introduction

In a previous report, Waddington (1960), details were given of a determination of the flux of primary cosmic ray protons in nuclear emulsions. In that report, in order to permit a correction to be made to the observed flux for the overlying atmosphere, an empirical, and perforce rather crude, model was proposed for the propagation of these protons through the atmosphere. At the same time it was pointed out that the capability now existed for obtaining experimental material for which these corrections would be much less important, thus reducing the dependence of flux values above the atmosphere on the somewhat doubtful validity of the proposed model. Such experimental material has now been obtained, and this paper reports the results obtained.

§ 2. Stack and Exposure Details

Ten 600 μ, 5 cm by 5 cm Ilford G5 stripped emulsions, total volume 15 cm^3, were exposed under less than 0·1 g/cm^2 of packing material, as a small stack. This exposure was made to the primary cosmic radiation on a high altitude balloon flown from Nacogdoches, Texas, on July 12, 1960. The balloon reached a residual pressure of $4{\cdot}5 \pm 0{\cdot}3$ mb and then floated for 10 hs 42 min before releasing the equipment near Lovington, New Mexico. The stack was mounted at least 15 m above any other massive part of the equipment so as to reduce the effects of splash albedo, and was rotated through 180° when the balloon reached a residual pressure of 11 mb, 30 min before reaching ceiling.

§ 3. Procedure

The processed emulsions from this stack were analysed in a manner closely similar to that adopted previously. A selected volume was searched for all those nuclear interactions in which three or more black or grey tracks were produced ($N_h \geqslant 3$). Each 'star' found in this manner was carefully examined for the presence of minimum tracks produced by fast singly charged particles. In the event that one or more of these lay within 60° of the vertical, the star was treated as one that could have been produced by a primary proton, the final decision being made from the configuration of all the minimum tracks associated with the star, as before. The flux, $J(X)$, was then given by

$$J(X) = \frac{\lambda dN}{V\pi t} \text{ particles/m}^2 \text{ ster sec}$$

where λ is the interaction mean free path of protons in emulsion, in metres, dN is the number of stars, V the volume, in m^3, and t the duration of the exposure in seconds.

§ 4 DETECTION INEFFICIENCIES

In order to determine dN, the number of stars found has to be corrected for various forms of detection inefficiency. The most serious of these is the correction for the loss of small stars, since these cannot be detected by area scanning. Previously a correction factor, Γ, of $1{\cdot}56 \pm 0{\cdot}07$† was calculated for those stars with $N_h \geqslant 5$ from the star size distribution found by Winzeler *et al.* (1960) for 648 stars produced by 6·2 BEV protons. Since then additional data has become available from 'along the track scanning' of artificially accelerated protons. Rajopadhye (1960) has examined 268 stars produced by 5·7 BEV protons, Aly (1960) 203 stars produced by 6·0 BEV protons, and the CERN group‡ 451 stars produced by 24 BEV protons. Γ derived from these experiments has a value of $1{\cdot}86 \pm 0{\cdot}06$, which appears significantly different from that derived previously. This difference does not appear to be a consequence of the high energy of some of these additional protons.

For want of a better procedure *all* this data has been used to calculate a mean Γ of $1{\cdot}74 \pm 0{\cdot}04$, but the possibility that this value, and thus the proton fluxes dependent on it, might have to be further changed, should be recognized. Figure 1 shows the summed size distribution from all these experiments, together with the size distribution obtained in this experiment. This latter distribution is normalized for those stars with $N_h \geqslant 5$.

These experiments also lead to values of λ, the interaction mean free path. Again taking all the data the mean value of λ is $37{\cdot}2 \pm 0{\cdot}8$ cm. It should be noticed that λ and Γ are not independent and that this will tend to minimize any errors in the quoted values.

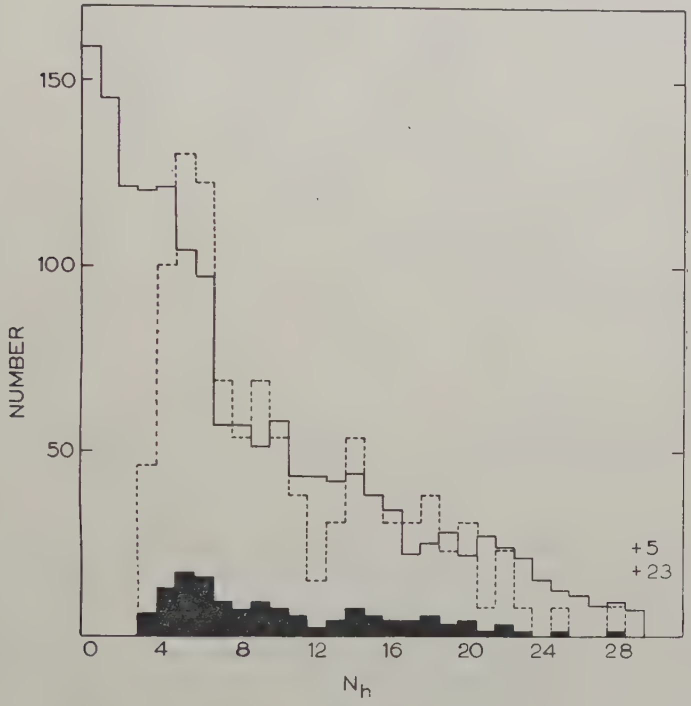
---

† This error is actually rather less than that given in the earlier work.

‡ J. C. Combe, W. M. Gibson, W. O. Lock, M. Schneeberger and G. Vanderhaeghe (private communications).

The other possible cause of detection inefficiency is the likelihood of failing to detect some of the tenuous tracks produced by the fast singly charged particles associated with each star. Such an effect would almost inevitably be a function of the zenith angle of the track. However, a study of the zenith angle distribution of the primary particles observed in this experiment showed no signs of any systematic missing, and so it has been assumed that no primary particles were missed because of a failure to observe them.

Fig. 1



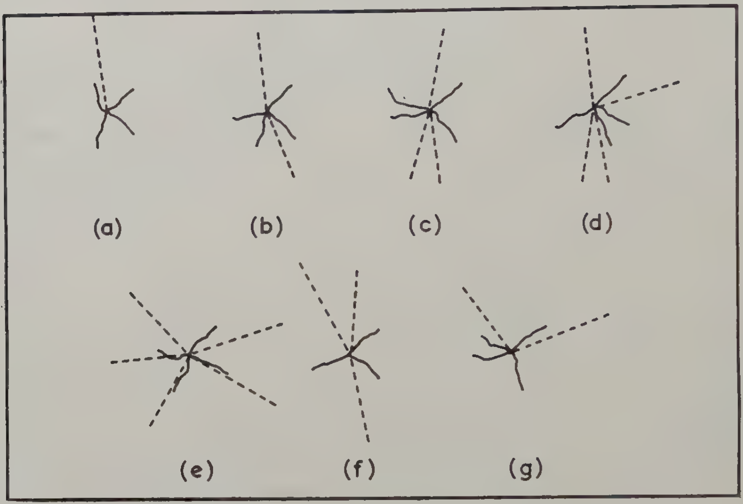
Total N_h distribution derived from experiments with artificially accelerated protons—solid line. N_h distribution observed in the experiment shown blocked in, while same distribution, normalized to previous distribution at $N_h \geqslant 5$—dashed line.

§ 5. IDENTIFICATION OF PRIMARY PARTICLES

In some, but not all stars, one of the minimum tracks has a much greater probability than any of the others of being that produced by the primary particle. In the previous paper the stars were separated into seven groups on a basis of the configuration of the minimum tracks. Figure 2 shows sketch drawings of typical members of each group. In groups (*a*), (*c*) and (*d*) the identity of the primary particle is well determined, while in (*b*) it is known with a high probability if group (*f*) is small. In groups (*e*) and (*g*) the primary is not well identified, and the stars must be given a fractional weight.

In this experiment, in a sample of 140 stars, the percentages in each group were as follows (the precentages found previously are given in brackets): (*a*) 14·3% (18%); (*b*) 17·9% (19%); (*c*) 45·0% (37%); (*d*) 8·6% (8%); (*e*) 10·7% (8%); (*f*) 2·1% (5%); and (*g*) 1·4% (5%). Any ambiguity in the identification of primary particles can thus be seen to be rather small.

Fig. 2



Sketch drawings of typical stars. Each star has $N_h=4$.

§ 6. Experimental Results

In scanning a volume of 0·99 cm³ 118 stars were found which had $N_h \geqslant 5$ and a definite or possible primary. Some of these were of group (*e*) or (*g*) and hence had to be given a weight of less than one. Thus this sample corresponded to 111·3 stars. This number was multiplied by Γ, equal to $1{\cdot}74 \pm 0{\cdot}04$, as derived in § 4, to give the true value of dN. Then, $J(X)$ was calculated from eqn. (1) to be

$$J(X)=602 \pm 60 \text{ particles/m}^1 \text{ ster sec}$$

when $t=642$ min, the time at ceiling altitude.

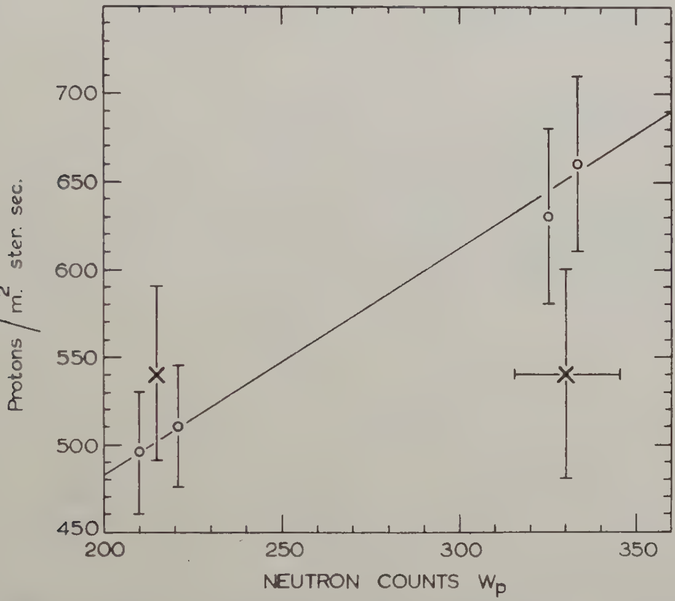
In order to obtain the flux at the top of the atmosphere, J_0, this value must be corrected; (i) for those particles which entered during the 30 min while the stack rose to the ceiling altitude after rotation; and (ii) for the effects of the overlying atmosphere. These corrections involve the use of the model derived previously. The version called model (*c*) was then assumed to be the most appropriate, and with this version correction (i) becomes a correction of 31 min in the time, so that the correction factor is 0·955, while correction (ii) gives a factor of 0·945. The total correction factor is thus 0·902, of which it should be noticed that (i) hardly depends at all on the model used. Thus

$$J_0=540 \pm 50 \text{ protons/m}^2 \text{ ster sec.}$$

This value is not too dependent on the model used to derive the correction to the top of the atmosphere, whereas previously, the total correction factor due to the ascent correction and that for the overlying atmosphere was 0·694, and the adjusted value obtained for J_0 of 540 ± 60 protons/m^2 ster sec would be seriously affected by inadequacies in the model.

These values of J_0 are compared with those derived by McDonald and Webber (1959), as quoted by Webber (1961), in fig. 3, which shows the proton fluxes plotted as a function of the Ottawa neutron monitor data, pertaining to the date of each flight. The value obtained in this experiment is in good agreement with the counter data, while even the previous result is not statistically significantly different, showing that it is unlikely that there are any serious errors in these data or in that of the counter determinations. However, both the previous values obtained at this geomagnetic cut-off, and one obtained at a lower cut-off, Stevenson and Waddington (1961) are appreciably, even if not significantly, lower than the counter results. Both these values are seriously dependent on the corrections, and it is therefore possible that the model used tends to overestimate these corrections.

Fig. 3



Proton fluxes obtained by counters, ○, and in emulsions, ×, plotted against the Ottawa neutron monitor data for each flight, W_p. W_p is expressed in the natural logarithm form $= \ln (N_p/1000) - 1{\cdot}500$ and N_p is the bi-hourly counting rate.

It should be emphasized, however, that this technique may not be applicable to the determination of fluxes of protons having minimum energies appreciably below those observed in this experiment. At lower energies the ambiguity between secondary and primary tracks may become more serious, and invalidate the technique, thus explaining the result obtained at a lower cut-off.

ACKNOWLEDGMENTS

The author is grateful to Professor C. F. Powell, F.R.S., for the facilities of his laboratory, and to Professor M. W. Friedlander of Washington University, St. Louis, for his invaluable association in the exposure and processing of these emulsions. This work was supported by a grant from the Department of Scientific and Industrial Research.

REFERENCES

ALY, H. H., 1960, Ph.D. Thesis, University of Bristol.
MCDONALD, F. B., and WEBBER, W. R., 1959, *Phys. Rev.*, **115,** 194.
RAJOPADHYE, V. Y., 1960, *Phil. Mag.*, **5,** 537.
STEVENSON, G., and WADDINGTON, C. J., 1961, *Phil. Mag.*, **6,** 517.
WADDINGTON, C. J., 1960, *Phil. Mag.*, **5,** 1105.
WEBBER, W. R., 1961, *Progress in Cosmic Rays*, Vol. 6 (to be published).
WINZELER, H., KLAIBER, B., KOCH, W., NIKOLIC, M., and SCHNEEBERGER, M., 1960, *Nuovo Cim.*, **17,** 8.

On the Annealing of Prismatic Dislocation Loops in Aluminium

By F. Kroupa†, J. Silcox‡ and M. J. Whelan
Cavendish Laboratory, Cambridge, England

[Received April 19, 1961]

Abstract

A recently proposed mechanism for coalescence of dislocation loops on annealing by a pipe diffusion mechanism is examined. The driving force for pipe diffusion arises from two sources. In the first, differences in concentrations of vacancies round the perimeter of a loop due to differences in the forces exerted by a neighbouring loop cause a net flow of vacancies tending to make the loops coalesce. In the second, vacancies, formed at a given loop, interact with the stress field of a neighbouring loop. The driving forces for both these processes seem to be of the same order of magnitude, and the dependence on separation of the loops is similar. Estimates of the driving force are made from the stress field of a prismatic loop. The driving force is very small owing to the rapid decrease of the stress field with distance from a prismatic loop. It is shown that the times of coalescence of typical neighbouring loops are much longer than those recently proposed. It is concluded that pipe diffusion cannot be the rate-controlling mechanism for annealing of loops in quenched aluminium.

In a recent paper Johnson (1960) has proposed that a mechanism of pipe diffusion around the perimeter of a dislocation loop (formed by vacancy aggregation in quenched aluminium) accounts more satisfactorily for the observed behaviour on annealing in the temperature range 125°c to 200°c than the bulk diffusion mechanism proposed by Vandervoort and Washburn (1960) and by Silcox and Whelan (1960). In the pipe diffusion mechanism, illustrated in fig. 1, a small dislocation loop is envisaged to move towards a larger loop by transport of material around the perimeter of the loop (a process called self-climb). It was suggested that if this occurs with a lower activation energy than that of self-diffusion, the process could account for the observed rate of growth of loops at 150°c more satisfactorily than the bulk diffusion mechanism. Experimental evidence for pipe diffusion has been given by Kroupa and Price (1961), who interpreted the sideways movement of a prismatic dislocation loop under the influence of the stress field of a neighbouring straight dislocation in a thin zinc foil in terms of a self-climb mechanism.

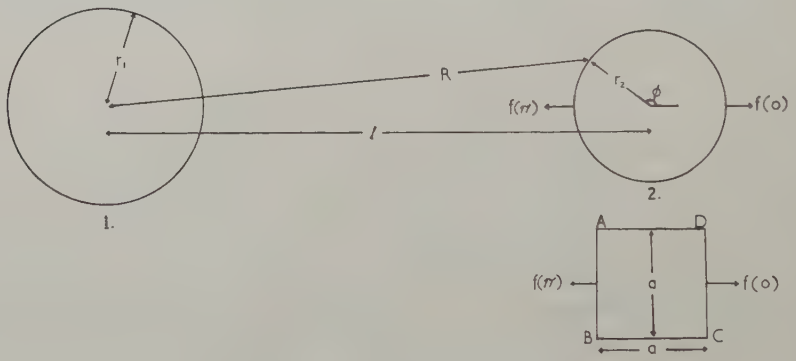
In considering the possibility of pipe diffusion, we first wish to point out that Johnson's numerical results are subject to error because he has neglected to consider the driving force in the calculation of diffusion rates both for bulk diffusion and for pipe diffusion. In the case of pipe diffusion, the driving force arises from two sources. The force on one loop due to the stress field of another loop gives rise to differences in concentration of vacancies around the loop (i.e. the energy of formation at the loop, U_f', is affected), resulting in a net flow of vacancies. Vacancies once created at a given loop may, however, interact with the elastic stress field of a second loop and accordingly the energy for migration, U_m' will be affected. We will consider these two effects independently for the sake of clarity and show that they are both of the same small order order of magnitude.

---

† On leave of absence from The Institute of Physics, Prague, Czechoslovakia.
‡ Now at Department of Engineering Physics, Cornell University, Ithaca, N.Y., U.S.A.

We consider the interaction between two loops 1 and 2 shown in fig. 1 in a simple cubic lattice of cell constant b. These loops are assumed to be lying in the same plane with Burgers vectors perpendicular to the plane. There is a force $f(\phi)$ per unit length on loop 2 due to the stress field of loop 1 which depends on the angular coordinate ϕ as well as on the separation l between the centres. If both loops are formed from vacancies the force is radial and acts in a direction to expand the loop. The driving force for self-climb will be determined essentially by the difference $\Delta f = f(\pi) - f(0)$ as the following considerations show.

Fig. 1



Nomenclature used in the text.

A detailed theory of the climb rate would discuss the change in shape of an initially circular loop. However, following Johnson, we will take for simplicity the case of a square loop 2 of side a as shown in fig. 1. We also assume that the force on the loop changes linearly from $f(0)$ at CD to $f(\pi)$ at AB. Imagine that the loop 2 behaves as a ‘ pipe ’ of cross-sectional area δ^2; where the energies of formation and movement of vacancies are U_f' and U_m' respectively. These energies are imagined to be lower than the corresponding energies of a vacancy in a perfect crystal. Then under the force $f(0)$, the segment CD will have an atmosphere of vacancies (Friedel 1956) of concentration $c_0 \exp(-f(0)b^2/kT) \simeq c_0(1-f(0)b^2/kT)$, if $f(0)b^2 \ll kT$ (satisfied for $l \gg r_1+r_2$ in fig. 1), where c_0 is the equilibrium concentration in the pipe under zero force. Similarly the segment AB will have an atmosphere of concentration $c_0\,(1-f(\pi)b^2/kT)$

The total flux of vacancies through the two pipes AD and BC in the sense CD→AB due to these concentration differences will be

$$2\delta^2 D\frac{c_0}{Va}\frac{[f(\pi)-f(0)]b^2}{kT}, \qquad . \; . \; . \; . \; . \; . \; . \; . \quad (1)$$

where V is the vacancy volume, and where $D\simeq z\nu_0 b^2 \exp(-U_m'/kT)$ is the diffusion coefficient for vacancy movement along the pipe. z is a coordination factor; ν_0 is the atomic vibration frequency $\sim 10^{13}$; c_0 will be $\simeq \exp(-U_f'/kT)$. Substituting for D, c_0 in eqn. (1), the flux may be written as

$$\frac{2\delta^2 z\nu_0 b^2}{Va}\exp\left(-\frac{U'}{kT}\right)\frac{\Delta f b^2}{kT}, \qquad . \; . \; . \; . \; . \; . \quad (2)$$

where $\Delta f=f(\pi)-f(0)$ and where $U'=U_f'+U_m'$ is the activation energy of pipe diffusion. The velocity of self-climb v is given by

$$\text{flux}=va/b^2.$$

Then

$$v=\frac{2\delta^2 z\nu_0 b^4}{Va^2}\exp\left(-\frac{U'}{kT}\right)\frac{\Delta f b^2}{kT}. \qquad . \; . \; . \; . \; . \; . \quad (3)$$

The correct values of δ^2 and z for pipe diffusion are rather obscure. However, we note that if we put $z\sim 2$, $\delta\sim b/2$, $V\sim b^3$, the first term in brackets in eqn. (3) is the same as the term derived from eqn. (7) of Johnson's paper. We then obtain

$$v=\left(\frac{\nu_0 b^3}{a^2}\right)\exp\left(-\frac{U'}{kT}\right)\frac{\Delta f b^2}{kT}. \qquad . \; . \; . \; . \; . \; . \quad (4)$$

Equation (4) differs from the corresponding velocity derived from Johnson's eqn. (7) by the factor $\Delta f b^2/kT$, which is the term taking account of the driving force omitted by Johnson.

If we take a more realistic value for the cross-sectional area of the pipe, say by including the five sites around any atom at the core of an edge dislocation, the first term in brackets in eqn. (4) may be increased by a factor ~ 20. However, such factors were not considered by Johnson, and eqn. (4) as it stands ought to be compared with the corresponding velocity derived from Johnson's eqn. (7).

The force $f(\phi)$ is $\tau_{zz}^{(1)}b_2$ where $\tau_{zz}^{(1)}$ is the stress due to loop 1 normal to the climb plane (i.e. normal to the paper in fig. 1) and b_2 is the Burgers vector of loop 2. Using the expressions for $\tau_{zz}^{(1)}$ given by Kroupa (1960) we obtain

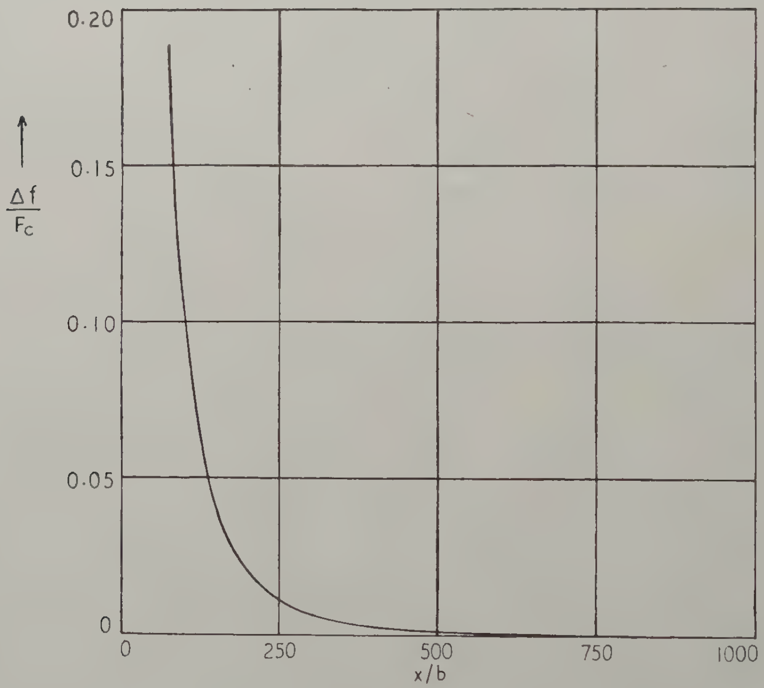
$$f(\phi)=\frac{Gb_1b_2}{\pi(1-\sigma)r_1}\frac{1}{\rho}\left[\left(1-\frac{1}{\rho^2}\right)^{-1}E\left(\frac{1}{\rho}\right)-K\left(\frac{1}{\rho}\right)\right], \qquad . \; . \; . \quad (5)$$

where $\rho=R/r_1$ (see fig. 1 for notation); $E(x)$ and $K(x)$ are the complete elliptic integrals (Jahnke and Emde 1933). G and σ are the shear modulus and Poisson's ratio respectively and b_1 is the Burgers vector of loop 1. If $b_1=b_2=b$, $\sigma=\frac{1}{3}$ and $r_1/b\sim 50$ to 100,

$$Gb_1b_2/\pi(1-\sigma)r_1\simeq Gb^2/2r_1\simeq F_c. \qquad . \; . \; . \; . \; . \; . \quad (6)$$

F_c is the force due to the line tension and curvature of loop 1 (Silcox and Whelan 1960); F_c provides a convenient unit of force. For l large compared with r_1, r_2, f/F_c varies as $1/\rho^3$ while $\Delta f/F_c$ will fall off as $1/\rho^4$. Figure 2 obtained from eqn. (5) shows how $\Delta f/F_c$ varies with the separation $x=l-(r_1+r_2)$ between the loops for $r_1=100b$, $r_2=50b$. The latter value is about the average radius of a typical loop distribution in aluminium quenched from 600°C. For this quenching temperature the average separation l between loops is about 1000 Å or 350 b, i.e. $x \approx 200b$. It is seen from fig. 2 that for this separation $\Delta f/F_c$ is about 2×10^{-2}. For aluminium, taking $G=2{\cdot}5\times10^{11}$ dynes cm^{-2}, $b=2{\cdot}86$ Å, $k=0{\cdot}86\times10^{-4}$ ev degree C^{-1}, $r_1 \sim 100b$, we have from eqn. (6), $F_c b^2/kT = Gb^4/2r_1kT \simeq 0{\cdot}5$ at 160°C. Therefore $\Delta f b^2/kT \simeq 10^{-2}$. The velocity of self-climb calculated from eqn. (4) is therefore smaller than that calculated by Johnson. For a separation $l \sim 3000$ Å, as was actually considered by Johnson, $\Delta f b^2/kT \sim 10^{-4}$.

Fig. 2



Plot of the function $\Delta f/F_c$ as a function of x/b.

To calculate the magnitude of annealing times, we derive the time t required for the separation l to decrease from l_1 to l_2. Following Johnson we ignore motion of the large loop. We take $a=r_1=100b$, and use the series expansion for $E(1/\rho)$ and $K(1/\rho)$ in eqn. (5). We then obtain

$$\Delta f/F_c = 3\pi/4\rho^4. \quad \quad (7)$$

Substituting (7) in eqn. (4), and putting $F_c b^2/kT \approx 0{\cdot}5$ at 160°c and $R \sim l$, we obtain a differential equation for l which may be integrated to give

$$t = 0{\cdot}85 \times 10^{-8}\tau(l_1{}^5 - l_2{}^5)/5l_1b^4. \qquad (8)$$

τ is the annealing time for coalescence of two loops of initial separation l_1 as given by Johnson's eqn. (7). Although the approximation (7) is valid only for large ρ, we should expect that (8) with $l_2 = 0$ will be a good approximation to the coalescence time because most of the time is taken when the loops are at a large separation. Putting $l_2 = 0$, $l_1/b = 1000$ in eqn. (8), we obtain $t \sim 2000\tau$. Annealing times by the pipe diffusion mechanism are therefore very much longer than those tabulated by Johnson. To obtain agreement with the observed times considered by Johnson the activation energy for pipe diffusion U' would have to be less than 0·5 ev, which is very low. Even if the factor mentioned above (taking account of an increased pipe diameter) is included the annealing times are still 100 times greater than those quoted by Johnson.

We now consider the second effect giving rise to a driving force, that due to the interaction of the vacancies once formed at loop 2 with the stress field of loop 1. Neglecting small changes in vacancy concentration in the pipe due to the effect discussed above, we assume an equilibrium concentration $\sim \exp(-U_f'/kT)$. The effect of the force $\mathbf{X}$ on a vacancy due to the stress field of loop 1 will be to impose a drift velocity $\mathbf{X}\, D/kT$. The resulting vacancy flux from CD to AB (fig. 1) in the direction of $\mathbf{X}$ is given by (cf. eqns. (1) and (2) with similar approximations for δ^2 and z)

$$\text{flux} = \frac{\nu_0 b}{a} \exp\left(-\frac{U'}{kT}\right)\frac{Xa}{kT}. \qquad . \quad . \quad . \quad . \quad . \quad (9)$$

This leads to an expression for the velocity v of the loop similar to eqn. (4), but with the driving force term replaced by Xa/kT. As in the previous case this factor may possibly be increased by ~ 20 to take account of a larger cross-sectional area of the pipe.

The nature of the interaction of a vacancy with a stress field is difficult to assess since the validity of the elastic model (i.e. a misfitting sphere) is open to doubt (Eshelby 1956). However, we will give a treatment following that given by Eshelby (1956) for the interaction of a point defect with a stress field. We consider the vacancy first as a dilation centre and the interaction energy, E_{int}, is then given as

$$E_{\text{int}} = -\Delta V p^{(1)} = \Delta V G r_1{}^2 b(1+\sigma)/2(1-\sigma)R^3, \quad . \quad . \quad . \quad (10)$$

where ΔV is the change in volume due to the defect and $p^{(1)}$, the hydrostatic stress component produced at the defect by loop 1, is given as $-Gr_1{}^2b(1+\sigma)/2(1-\sigma)R^3$ (Kroupa 1960) for large R. The force $\mathbf{X}$ on the vacancy is therefore

$$\mathbf{X} = -\operatorname{grad} E_{\text{int}} = 3\Delta V G r_1{}^2 b(1+\sigma)\mathbf{R}/2(1-\sigma)R^5.$$

If we now take $\Delta V = -\epsilon b^3/\sqrt{2}$ where ϵ is a measure of the relaxation around a vacancy ($\sim 0{\cdot}2$; see e.g. Benneman and Tewordt 1960), this expression may be put in the form (with $\sigma = \frac{1}{3}$)

$$X = 2Gb^2(r_1/b)^2/5(R/b)^4. \qquad (11)$$

Equation (11) shows that X has the same dependence on R as Δf considered previously (see eqn. (7)). In fact inspection of (11) shows that for large R, with $a = r_1$,

$$Xa/kT \approx \Delta f b^2/3kT.$$

Therefore the contribution of this mechanism to pipe diffusion is slightly smaller than but of the same order of magnitude as that of the previous mechanism. Since the dependence on R is similar in the two cases, both mechanisms are of equal importance. Vacancies may also interact with a stress field since they are relaxation centres (i.e. $E_{\text{int}} \sim e \times$ atomic volume where e is the density of elastic energy due to loop 1). This interaction is much smaller than the dilatation effect (with the value of ϵ assumed here) and is neglected.

In considering the mechanisms of bulk diffusion, Johnson again neglected the driving force, which in this case is due to the line tension force F_c and the chemical stress as discussed by Silcox and Whelan (1960). If one neglects the chemical stress initially, inclusion of the driving force due to F_c does not appreciably affect the magnitude of annealing times since the driving force factor $[\exp(F_c b^2/kT) - 1]$ (Silcox and Whelan 1960) is of order unity for loops of average radius. Annealing times calculated by Johnson in this case are only one order of magnitude greater than those derived by Silcox and Whelan for loops of average radius in thin foils. This factor of 10 appears to arise from omission of a coordination factor $z \sim 11$ included by the latter authors. If this factor is taken into account, then the times calculated by Johnson for bulk diffusion at 150°C (i.e. now 2×10^3 sec) are only a factor of 3 greater than the annealing times used by Vandervoort and Washburn ($\sim 6 \times 10^2$ sec). If we now note that inclusion of the chemical stress can alter the sign of the driving force for the larger loops with values again of the order of unity, this suggests that bulk diffusion cannot be ruled out on the grounds that the process is too slow.

In addition to these points, there appear to be several experimental facts suggesting that bulk diffusion and not the self-climb mechanism is the rate-controlling process. These are as follows.

(*a*) Federighi (1959) has studied the annealing of the electrical resistivity in quenched aluminium in the range of temperatures 130 to 170°C, a range covering the experiments of Vandervoort and Washburn. The most reasonable interpretation of these experiments is that the decrease in electrical resistivity is due to changes in the dislocation loop distribution. Throughout the entire range of annealing, 130–170°C, the curves were consistent with a unique activation energy of $\sim 1{\cdot}3$ ev.

(*b*) The self-climb process has not been observed in thin foils of aluminium. It has been pointed out earlier that the stress system of a prismatic loop falls off very rapidly with distance from the centre ($\sim 1/\rho^3$). The stress field of a loop of radius r_1 at the centre of the foil will be very similar to that in bulk material if the thickness of the foil is, for example $\gtrsim 5r_1$, since the stresses at the surface of the foil are then very small. In fact for a loop of average radius (~ 140 Å) at the centre of a foil ~ 2000 Å thick, this condition is easily satisfied. It is expected that occasionally such a loop would have been observed. In addition, for a loop on an inclined plane we may expect image forces of a similar nature to those between two loops. We would therefore expect to see self-climb, either two loops moving towards each other or one loop climbing towards the surface. This has not been observed.

(*c*) We finally consider the consistency of the experimental observations of Vandervoort and Washburn and of Silcox and Whelan. There appears to be a difference in annealing times of a factor ~ 2. If we assume a relationship between time and temperature of the form $t = t_3 \exp(E/kT)$ (which is true in both the pipe diffusion model and the bulk diffusion model), this time difference may be transformed into a temperature difference of ~ -7°c (using $E = 1{\cdot}3$ ev) at temperatures ~ 180°c. Taking this into account, the relative fraction of defects annealed out after 10 min at various temperatures is given in the following table.

Temperature	125°c†	150°c†	173°c‡	183°c‡	193°c‡	203°c‡
Fraction of loops annealed out	0·30	0·50	0·75	0·90	0·99	1·00

† Taken from Vandervoort and Washburn (1960).
‡ Taken from Silcox and Whelan (1960) and corrected by −7°c for the shorter annealing time employed (assuming an activation energy of 1·3 ev).

These figures form a reasonably continuous set and, in the absence of more detailed loop density counts, can be considered consistent.

While of course we realize that pipe diffusion is a feasible mechanism of coalescence of loops which are very close together ($l \sim r_1 + r_2$) when the forces are very large, it is difficult to see how the process can be rate-controlling for the usual separations and sizes of loops observed in quenched aluminium. We conclude that the bulk diffusion mechanism is still the most satisfactory interpretation of the annealing of such loops.

ACKNOWLEDGMENTS

We are indebted to Professor N. F. Mott, F.R.S., and to Dr. W. H. Taylor for their continued interest. Thanks are due to Dr. P. B. Hirsch for useful discussions. Acknowledgments for financial support are due to the British Council (F.K.), to D.S.I.R. (J.S.), and to the Royal Society Mr. and Mrs. John Jaffé Donation (M.J.W.).

REFERENCES

BENNEMAN, K. H., and TEWORDT, L., 1960, *Z. Naturf.* A, **15,** 772.
ESHELBY, J. D., 1956, *Solid State Physics*, 3 (New York: Academic Press), p. 79.
FEDERIGHI, T., 1959, *Phil. Mag.*, **4,** 502.
FRIEDEL, J., 1956, *Les Dislocations* (Paris: Gauthier–Villars), p. 66.
JAHNKE, E., and EMDE, F., 1933, *Tables of Functions* (Leipzig: Teubner), p. 46.
JOHNSON, C. A., 1960, *Phil. Mag.*, **5,** 1255.
KROUPA, F., 1960, *Czech. J. Phys.* B, **10,** 284.
KROUPA, F., and PRICE, P. B., 1961, *Phil. Mag.*, **6,** 243.
SILCOX, J., and WHELAN, M. J., 1960, *Phil. Mag.*, **5,** 1.
VANDERVOORT, R., and WASHBURN, J., 1940, *Phil. Mag.*, **5,** 24.

The Effect of Fatigue Stressing on the Diffusion Rate of Zinc in Aluminium

By S. PEARSON, A. J. BOARD and C. WHEELER
Metallurgy and Physics Department, Royal Aircraft Establishment,
Farnborough, Hants

[Received March 17, 1961]

ABSTRACT

Experiments have been made to find the effect of fatigue stressing on the diffusion rate of zinc in aluminium and in an aluminium–20% zinc alloy. Diffusion measurements were made on unstressed specimens and on specimens which were being fatigued in the temperature range 300°C–375°C. The results show that only small changes, at most, were produced in the diffusion rates despite the very high plastic strain rates (approximately 144 per hour) given by the fatigue stressing.

---

§ 1. INTRODUCTION

RECENT work on the fatigue of metals has shown that metallurgical changes can occur on the micro scale during fatigue stressing at ordinary temperatures which would not be expected to occur at the same temperature in the absence of stress. These effects have been discussed in a recent review of the fatigue of metals by Thompson and Wadsworth (1958) and they seem to be consistent with large local increases in diffusion rates. No quantitative macroscopic diffusion measurements have been reported on metals undergoing fatigue but measurements have been made of diffusion rates during creep which might be expected to produce similar effects to fatigue. Buffington and Cohen (1952) measured the self-diffusion rate of α-iron during creep under compressive stress at 890°C. They found no increase in diffusion rate under stress until flow occurred when the diffusion rate increased linearly with rate of strain, the diffusion coefficient increasing to twelve times its original value for a strain rate of $0{\cdot}25\ \mathrm{hr}^{-1}$. More recent work at the same laboratory by Ujiije *et al.* (1958) has cast doubt on the values, however, and it now seems that the increase in diffusion rate may be much smaller than originally reported. Lee and Maddin (1959) measured the self-diffusion rates of silver during torsional creep and they report apparent increases in diffusion rates by factors as large as one hundred for strain rates of $0{\cdot}86\ \mathrm{hr}^{-1}$. On the other hand, Darby *et al.* (1959) measured the self-diffusion rates of silver during plastic deformation in extension and compression and even though the strain rates they used were in the range $0{\cdot}04$ to $1\ \mathrm{hr}^{-1}$ they obtained no significant increases in diffusion rates.

We have measured the diffusion rates of zinc in aluminium and in an aluminium–20% zinc alloy for specimens subjected to fatigue stressing in the temperature range 300–375°C and compared the results with measurements made on unstressed specimens over the same temperature range.

§ 2. EXPERIMENTAL

2.1. *Materials*

The materials used for the investigation were super purity aluminium (99·99%) and an alloy containing 20 weight per cent of zinc made from super purity aluminium and high purity zinc.

2.2. *Diffusion Measurements*

The fatigue specimens were 5 mm thick and had a centre section 9·5 mm wide and 16 mm long, while the unstressed specimens were circular discs, of the same thickness, and 11 mm diameter. All specimens were machined to size and then annealed for 16 hours, the annealing temperature being 550°C for the pure aluminium and 450°C for the alloy. After this treatment the grain size was about 1·2 mm for the pure aluminium and 1·0 mm for the alloy. A layer of radioactive ^{65}Zn approximately 8 μ thick was plated onto the surface, the zincate process being used to ensure that there was no oxide barrier between the aluminium and the zinc. For the fatigue specimens the radioactive layer was confined to an 8·3 mm diameter circle in the centre of the parallel portion by means of a suitable stainless steel jig. A layer of aluminium approximately 1 μ thick was then deposited over the zinc to reduce evaporation and oxidation during the experiments.

After the diffusion anneal, which was similar for stressed or unstressed specimens, the centre plated portions were cut out from the fatigue specimens and turned into discs. The unplated surfaces of all the discs, stressed or unstressed, were machined to remove the effects of any diffusion along the surface. The discs were then mounted in Wood's metal and successive layers were removed by grinding on wet carborundum powder. The reduction in thickness was measured with a micrometer reading to 1 μ and at each stage the residual activity was measured under standard conditions on a scintillation counter. For comparison some specimens were sectioned by the usual lathe sectioning technique and good agreement was obtained between the two methods.

Preliminary experiments on unstressed specimens had shown that diffusion occurred at the same rate in air or in vacuum so for the main investigation all experiments were made in air as this simplified the fatigue apparatus.

2.3. *Fatigue Experiments*

Two machines were used for stressing the fatigue specimens. One was a modified Haigh electromagnetic machine which applied an alternating tension compression stress at 2000 cycles per minute and the other a specially designed cantilever machine which applied a tension compression strain also at 2000 cycles per minute. In the Haigh machine the stress was uniform over the full cross section of the specimen whereas in the cantilever machine the strain varied over the specimen cross section but over the maximum diffusion distance the change in strain was less than 2% which was considered to be sufficiently uniform for the experiments.

The experiments usually lasted for about 30 hours and were in the temperature range 300–375°c, the upper limit being set by the excessive roughening of the surface of the pure aluminium specimens and the very low fatigue strength of the alloy above this temperature. These diffusion conditions gave measurable penetrations of the zinc into the aluminium of about 60–350 μ.

The Haigh machine was used mostly for fatigue experiments on the pure metal, although a few experiments were made on the alloy. Most experiments were made at stresses of about 80–90% of those to cause failure in 30 hours, although some experiments were made at higher stresses and the specimens failed before the 30 hours was completed. Failure was always transcrystalline for the pure metal and intercrystalline for the alloy. A reasonable proportion of the failures occurred through the plated region and these specimens were discarded but where failure occurred outside this region the specimens were sectioned in the usual way. The surface of the pure aluminium specimens roughened during fatigue, in some cases so severely that the specimens were unsuitable for sectioning and, although most of the specimens could be sectioned, the accuracy of the diffusion measurements was reduced. The surface of the alloy specimens was not roughened, about 30% of the grains showing coarse slip and the remainder fine slip only. There was also evidence of considerable deformation at the grain boundaries and of grain boundary migration.

Experiments were made on the pure aluminium and the alloy in the cantilever machine, the maximum strain at the surface of the specimen being $4\cdot5\times10^{\ 4}$ in all cases. With this strain the surface of the pure aluminium was roughened and in some cases, usually at the lower temperatures, failure occurred before the completion of 30 hours. The appearance of the alloy specimens was similar to those fatigued in the Haigh machine and most specimens fatigued at temperatures greater than 325°c failed before the completion of 30 hours.

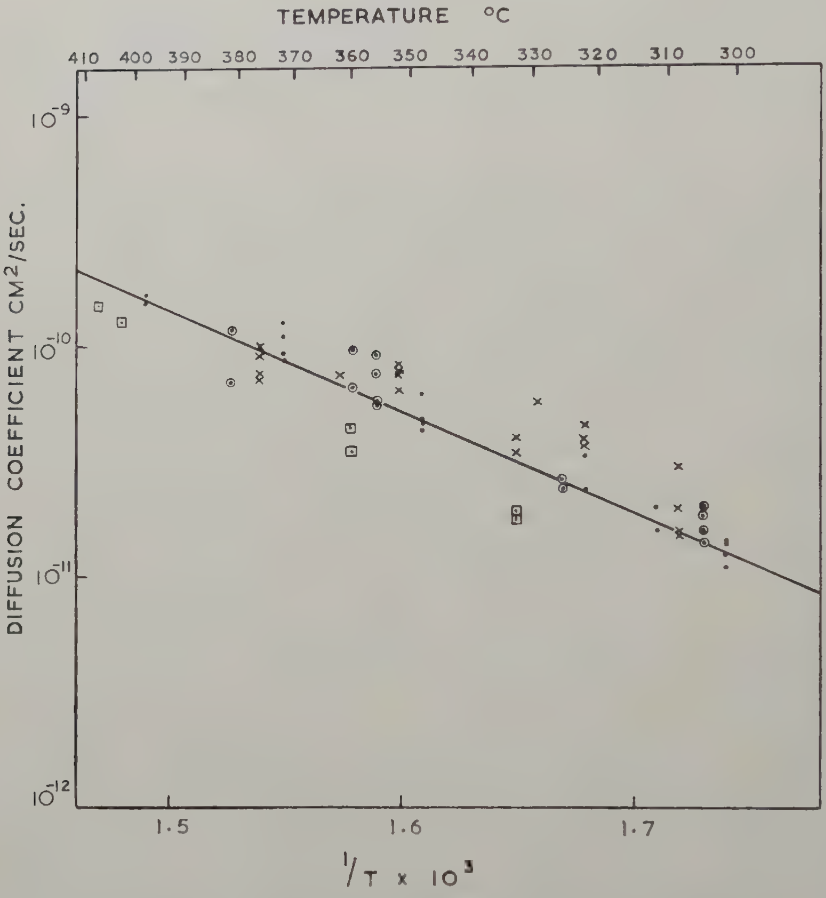
§ 3. RESULTS

The results of the experiments are given for the pure aluminium in fig. 1 and for the alloy in fig. 2.

§ 4. DISCUSSION

From fig. 1 the effect on the diffusion rate of zinc in the aluminium was an increase of about 25%, although statistical analysis shows that the increase is within the random experimental error. From fig. 2 the effect of fatigue on the diffusion rate in the alloy was a decrease of about 35%. It is difficult to see why this decrease should occur and no explanation for it is offered but it is outside the random error and no systematic error could be found to account for it.

Fig. 1

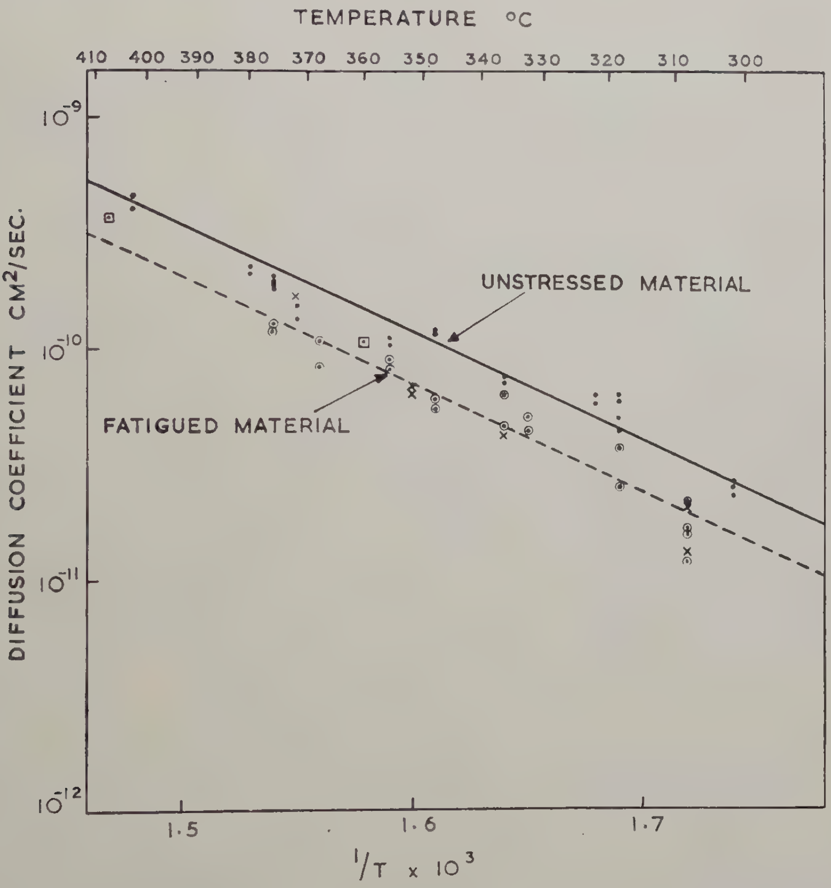


Effect of fatigue stressing on diffusion of zinc in aluminium.

● Unstressed material.
⊙ Fatigued cantilever machine.
× Fatigued Haigh machine.
⊡ Results obtained by Hilliard *et al.* for unstressed material.

Recent measurements of the diffusion rate of zinc in unstressed specimens of aluminium and of aluminium zinc alloys have been made by Hilliard *et al.* (1959). Their results for pure aluminium are shown in fig. 1 and for a 20% zinc alloy in fig. 2. It will be seen that in each case they are slightly lower than the results we obtained and it seems that our results may have been influenced by grain boundary diffusion at the low temperatures used, or by the concentration gradient in our experiments. However, these effects are evidently small and do not seriously affect the direct comparison between the unstressed and the fatigued specimens.

Fig. 2



Effect of fatigue stressing on diffusion of zinc in an aluminium–20% Zinc alloy.

- ● Unstressed material.
- ⊙ Fatigued cantilever machine.
- × Fatigued Haigh machine.
- ⊡ Results obtained by Hilliard *et al.* for unstressed material.

An estimate of the plastic strain rate can be made by comparing the results obtained with the two machines. For pure aluminium at 300°c a peak stress of 70 kg/cm² produced failure in about the same number of cycles as a peak strain of 0·00045 so that allowing for the reduction in modulus with temperature the elastic strain was $\simeq 0{\cdot}00015$ giving a peak plastic strain of $\simeq 0{\cdot}0003$ per cycle and a total plastic strain rate of approximately 144 hr^{-1}. The maximum rate of strain in the published creep experiments was about 1·0 hr^{-1}, so that even if allowance is made for considerable dislocation reversal, rather than the creation of new dislocations, in the fatigue case it seems that if the large increases in diffusion rates during creep are correct, then correspondingly large increases should have been obtained in these experiments.

Where increases in diffusion rates have been obtained in the creep experiments they have usually been attributed to the large number of excess vacancies which would be continuously produced as the metal was strained. However, Lomer (1958) has shown that a non-equilibrium concentration of vacancies produced by irradiation damage at the diffusion temperature is unlikely to lead to a measurable increase in diffusion over macroscopic distances, although it could lead to increased diffusion over very small distances at low temperatures. Unless, therefore, vacancies can be produced at a much faster rate by plastic deformation than by the irradiation damage considered by Lomer the same conclusion will apply. An upper limit to the number of excess vacancies can be obtained by assuming that all the strain energy dissipated is used to form vacancies. In these experiments for pure aluminium at 300°c the energy dissipated, obtained from the product of stress and plastic strain per cycle, was $\simeq 2\times10^{6}$ ergs/cm^3/sec, so that taking a formation energy of 0·8 ev for vacancies in aluminium the maximum rate of production of vacancies was $\simeq 3\times10^{-5}$ per atom per second. Following Lomer (his eqn. (6.1)) the diffusion distance produced by the excess vacancies (the distance in which the concentration falls to $1/e$ of its value at the surface)

$$L=a\sqrt{(\tfrac{2}{3}rtn)}$$

where

a = atomic distance ($\simeq 3$ Å for aluminium),
r = rate of production of vacancies per sec,
t = time in secs,
n = number of jumps made by a vacancy before it is annihilated.

Lomer takes $n\simeq10^{6}$ and this seems to be a reasonable figure for heavily worked aluminium, so that for a production rate of 3×10^{-5} per sec and a diffusion time of 10^{5} sec (approximately 30 hrs) the diffusion distance is $\simeq 0{\cdot}5\,\mu$ which is much less than the distance produced thermally in unfatigued specimens ($\simeq 25\,\mu$). It seems, therefore, that even for the large strain rates given by fatigue stressing the excess vacancies produced are unlikely to cause a measureable increase in the macroscopic diffusion rate at these temperatures. They could, however, lead to increased diffusion over very small distances at lower temperatures so the results are not inconsistent with the apparent local diffusion effects which have been observed under the microscope.

§ 5. Conclusion

It is concluded from these experiments that despite the very high strain rates given by the fatigue stressing there was no significant increase in the diffusion rates. The results seem, therefore, to be at variance with those investigations which have given large increases in diffusion rates during creep deformation but they are in agreement with the results of Darby, Tomizuka and Baluffi.

Acknowledgments

We are grateful to Dr. N. J. Wadsworth for helpful discussions and advice and to Mr. W. Stalker who remelted and formed some of the pure aluminium and made up the aluminium–20% Zinc alloy.

References

Buffington, F. S., and Cohen, M., 1952, *Trans. Amer. Inst. min. (metall.) Engrs*, **194,** 859.

Darby, J. B., Tomizuka, C. T., and Baluffi, R. W., 1959, *J. appl. Phys.*, **30,** 104.

Hilliard, J. E., Averbach, B. L., and Cohen, M., 1959, *Acta Met.*, **7,** 86.

Lee, C. H., and Maddin, R., 1959, *Trans. Amer. Inst. min. (metall.) Engrs*, **215,** 397.

Lomer, W. M., 1958, *Vacancies and other point defects in metals and alloys* (Inst. of Metals Monograph No. 23).

Thompson, N., and Wadsworth, N. J., 1958, *Advanc. Phys.*, **7,** 72.

Ujiije, N., Averbach, B. L., Cohen, M., and Griffiths, V., 1958, *Acta Met.*, **6,** 68.

Measurement of Low-energy Beta-ray Spectra using a High-pressure Proportional Counter

By Paul E. Shearin and Robert L. Hoover
The University of North Carolina, Chapel Hill, N.C.

[Received March 23, 1961]

Abstract

In order to eliminate self-absorption in solid and liquid sources in the study of low-energy beta-particle emitters radioactive gas is introduced in a proportional counter. The low-energy spectrum of ^{14}C is measured and the experimental data agree with the theoretical plot from 0·5 kev upward to 7 kev within statistical limits. The cathode of the counter is made of wires and is surrounded by a ring of nine anti-coincidence guard counters. The main-counter anode is unequally divided to minimize end effects.

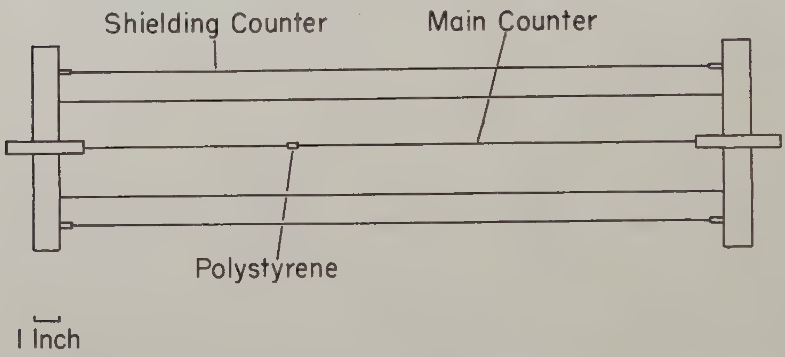
§ 1. Introduction

The results of an investigation in the low-energy portion of the ^{14}C beta-ray spectrum are presented. Beta-ray spectra have been extensively studied in recent years from both theoretical and experimental points of view. In general, theory and experiment are in agreement as regards the main features of the spectra. There is, however, a paucity of data for that portion of the spectra which lies below 7 kev. At the same time it is in this very region where coulomb effects and the effects of screening by orbital electrons should be observed. The lack of data in this low-energy region is due in part to the self-absorption of beta-rays in the source when the radioactive source is either solid or liquid. It is also in this energy region that electronic noise becomes most troublesome. Accordingly, an investigation of the low-energy portion of the ^{14}C beta spectrum, an allowed spectrum, was undertaken with an especially designed proportional counter into which the beta emitter was introduced in gaseous form. Self-absorption was thus avoided. Special effort was made to reduce circuit noise. The spectrum of ^{14}C has been accurately measured for energies greater than 7 kev, and in this region theory and experiment are in good agreement. In this study, therefore, primary emphasis was placed on improving the measurements in the portion of the spectrum below 7 kev, these having been extended down to 0·5 kev.

§ 2. Apparatus

The proportional counter is designed to reduce spurious counts in the low-energy region originating from high-energy beta-particles and background sources such as radioactive contamination and cosmic rays. The effects of low counter amplification near the ends of the counter are reduced by dividing the main-counter anode into two unequal sections and subtracting the data of the short section from the long section as introduced by Angus *et al.* (1949). Wall effects in which energetic beta-particles leave the sensitive region of the counter before they lose all their energy in ionization are reduced by the combination of a wire cathode surrounded by a ring of guard counters electrically connected in anti-coincidence with the main counter as introduced by Moljk *et al.* (1957). The small area of the wire cathode reduces the spurious background originating in the cathode material plus secondaries produced there from cosmic rays and the guard counters shield the main counter from primary cosmic rays and their secondaries. The instrument is further shielded from cosmic rays by means of 11 in. of iron placed above the counter.

Fig. 1



General arrangement of the side view of the proportional counter. The long anode is 15 in. and the short anode is 9 in.

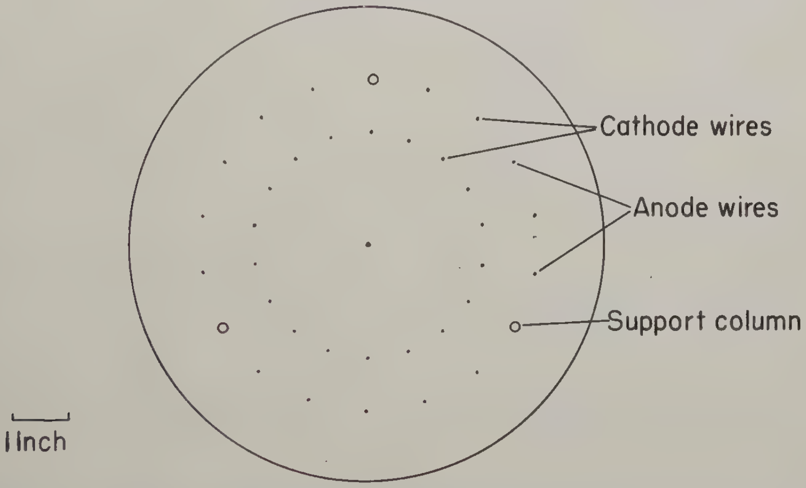
A side view of the counter is shown in fig. 1. The longer anode is 15 in. and the shorter anode is 9 in. both being 0·001 in. tungsten wires. These wires are stretched between polystyrene discs and are held in place by phosphor-bronze springs that tie onto steel connectors wedged into the polystyrene discs. The two anode sections are joined by a polystyrene insulator that is $\frac{1}{2}$ in. and $\frac{1}{16}$ in. in diameter. The polystyrene discs are held together by three steel rods $\frac{1}{4}$ in. in diameter.

The end view in fig. 2 shows the guard counters using the common cathode of the main counter and the steel enclosure. The unit composed of the main counter and the guard counters is inserted in a steel pipe 8 in. in diameter which becomes the exterior boundary for the guard counters. The main-counter cathode is 4 in. in diameter and the effective diameter of each guard counter is 2 in. The cathode is made of 18 steel wires, 0·007 in. in diameter. The wires are stretched between the polystyrene discs and are held taut by springs wound in the wire itself. The springs are covered by hollow steel cylinders forced into the polystyrene. All the cathode wires are connected to a circular terminal strip attached to each polystyrene disc.

Because the guard counters are connected electrically in anti-coincidence with the main counter, accurate measurements may be made of ionizing events which take place completely within the volume of the main counter by rejecting events which originate in the main counter and end beyond the cathode, or those events which originate beyond the cathode and end within the counter volume. Thus, the low energy portion of a beta-particle spectrum is not distorted by false counts originating from high-energy beta-particles abruptly stopped by a solid cathode.

Since guard counters are ineffective in reducing the gamma-ray background associated with cosmic rays the entire counter is covered with 11 in. of steel. The apparatus is located in a basement room where there is about 150 g/cm^2 of building material above; thus there is a total of about 367 g/cm^2 of absorber above the counter. Lead was not used for the shield because it contains so many radioactive impurities that the benefits obtained would be lost by generating background from the lead itself (Bearden 1933).

Fig. 2

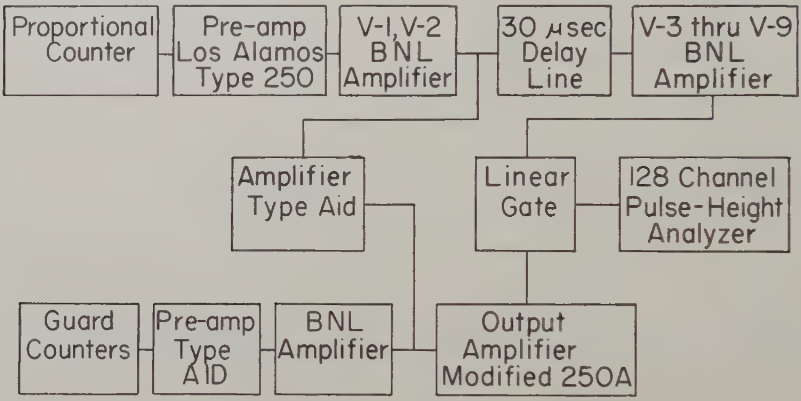


General arrangement of the end view of the proportional counter. The main counter diameter is 4 in.

The counter gas is 100 lb/in^2 gauge pressure made up of 80% argon and 20% methane. In all of the experiments the counter gas was cleaned for at least three days by passing the gas over calcium and magnesium heated to 600°C. The main-counter anode is set at 3900, 4025, 4200 and 4325 v in order to observe overlapping energy portions of the ^{14}C spectrum. The guard-counter voltage is set at 4200 v for all the ^{14}C runs.

Steps were taken to substantially reduce hum and noise so that pulses corresponding to the L-capture line of ^{37}A were more than 12 dB above the noise level. A block diagram of the electrical system is shown in fig. 3. Pulses in the main-counter amplifier are delayed 30 μsec so that the guard-counter amplifier will send initially simultaneous pulses ahead in the circuitry to block the flow of the main-amplifier pulses. The same voltage is maintained on both anode wires of the main counter; however, the pre-amplifier is attached to only one end at a time. An upper limit with regard to pulse height is selected by the gate circuit in addition to its anti-coincidence action in response to the guard counters. The output of the pulse-height-selector univibrator circuit in the A1D amplifier is connected to disable the gate circuit. If main-counter pulses pass through the gate, they then proceed to the pulse-height analyser, a 128 channel unit.

Fig. 3



Block diagram of the electronic system.

Anode voltage is derived from a series of 300 v dry batteries. The high voltage is by passed with a 2 μF, 5000 v capacitor and the dry batteries are coupled through a 100 kohm resistor to the capacitor.

The pre-amplifier was built after the design of the Los Alamos Laboratories (Johnstone 1955) and uses two Western Electric golden-grid tubes. The filament voltage is obtained from a regulated 300 v, 500 mA power supply. The main-counter and guard-counter amplifiers were built after the design of the Brookhaven National Laboratory (Chase and Higinbotham 1952). The amplifiers are powered with d.c. voltage on the filaments and regulated voltage on the plates. The chassis bottoms are covered with a sheet-copper shield rimmed with a metal mesh to make complete electrical contact. The linear gate circuit uses recently developed tubes and d.c. stabilization is provided for the difference amplifier stage. Second base-line crossings are produced in the low-energy portion of the ^{14}C spectrum by pulses correspondingly near the end-point energy. A gate circuit is adjusted to reject these high-amplitude pulses so that the second base-line excursions are prevented from being counted with the low-energy ^{14}C data. An all transistorized pulse-height analyser is used. The analyser has 128 channels and it may be set in an 'add' or 'subtract' mode. Thus the pulses from the long counter are added to the memory and pulses from the short counter are subtracted from the memory.

Radioactive carbon was prepared in $^{14}CO_2$ form from barium carbonate. The total activity of the $^{14}CO_2$ was limited to fewer than 60 000 disintegrations per minute in the guard-counter volume, so that the total dead time resulting from the anti-coincidence gate circuit is less than 0·4 sec in 1 sec. Overload pulses originating in the main amplifier increase the dead time to around 0·5 sec in 1 sec.

§ 3. Calibration

In order to determine the characteristics of the proportional counter a thorough study of ^{37}A was made prior to the ^{14}C investigation. The short counter has a calculated capacity of 1·64 $\mu\mu$F less than the long counter. For this preliminary investigation a 1·64 $\mu\mu$F capacitor was connected across the input anode terminal of the short counter. With this change there was only 1% increase in the half-width of the K-capture line as compared with the data taken with no capacity correction. Therefore, the correction capacitor was not used in further experiments.

^{37}A was then mixed with ^{14}C for calibration. The counter voltages were chosen on the basis of these ^{37}A and ^{14}C runs and the pulse amplifier was set at a constant amplification and only this gain or a multiple of ten thereof was used for the study. The ^{14}C and ^{37}A calibration runs simulated the actual conditions existing in the system while studying radioactive carbon alone. The normal rate of ^{14}C was introduced but only two effective disintegrations per minute of ^{37}A were introduced. Thus a small K-capture peak could be observed to build up on top of the carbon spectrum. 2·77 kev is used as the calibration energy which accounts for the average ionization potential of the counter gas. The pulse amplitude points for the 2·8 kev line corresponding to the counter voltage used in the carbon study lie near a straight line when plotted on a logarithmic scale.

The pulse heights increase in value with respect to increasing channel numbers. The curve of the channel numbers versus pulse heights for the complete system, including the proportional counter and the gate circuit, rises in a monotonic fashion, although the rise is not linear. Therefore, two calibration steps in the data reduction process are necessary. The response or channel number must be converted to a signal or pulse height and the actual number of accumulated pulses in each channel must be weighted to account for system non-linearity.

All the data reduction depends on the calibration curves prepared by feeding artificial pulses of known height from an AEC Model 506 pulse generator into the system and then reading the pulse-height analyser. The response versus signal calibration is a simple 1 : 1 correspondence of channel number to most favourable pulse height. In all amplifier configurations the pulse generator never spilled more than 10% of its pulses into side channels when the generator was adjusted to give a peak number of pulses in the middle channel. This figure represents the worst condition.

The weighting factor is a unique function and it is derived from the signal versus response curve. If $S=S(R)$ gives the 1 : 1 relation between signal and response where $g(S)\,dS$ is the number of pulses between S and $S+dS$, and where $G(R)\,dR$ is the number of signals between R and $R+dR$, then

$$g(S)=\frac{dR}{dS}G[R(S)]\simeq\frac{\Delta R}{\Delta S}\,.\,G[R(S)].$$

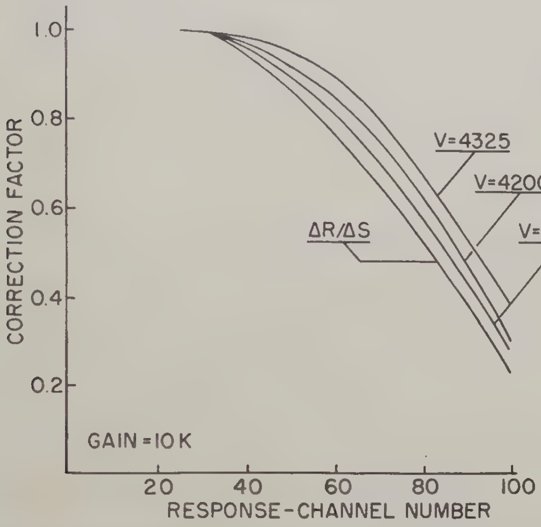
$\Delta R/\Delta S$ is obtained from the signal versus response curve.

For high counter gains the near maximum energy beta-particles produced an overloading of the amplifier, but this was corrected by appropriate circuits. However, the overall amplifier system changes its characteristics when an isotope is put in the counter, even though the circuits eliminate the strong overload pulse and its second base-line crossing. When the counter is connected to high voltage and only background pulses appear, the pulses from the artificial pulse generator group in a single channel. However, when ^{14}C is put into the counter this grouping is smeared out. The peak of the pulses does not change channels, but some counts are added to lower channels, whereas no counts are added to higher channels. Evidently, this condition arises from amplifier overload.

In order to correct for this effect the calibration curve is prepared in a circuitous way. Artificial pulses from the pulser are generated with a fixed pulse height and fed into the pulse-height analyser along with radioactive carbon pulses. After this is done for a measured time the pulse height is increased by a known increment and new pulses are added to the old pulses. This procedure is repeated so that all 128 channels are covered in 1000 equally divided steps of 15 sec duration each. The generator is then turned off, but the proportional counter is left on and the pulses from radioactive carbon are subtracted from the pulse-height memory for a time equal to the total generator running time. The data that are left approximate the configuration that random pulses of equal probability in all signal values from zero to some large value should create. The weighting factor is then obtained from these data by dividing the number of pulses in each channel into the smallest number in the group. The actual weighting factor is taken from a smooth curve drawn through these points.

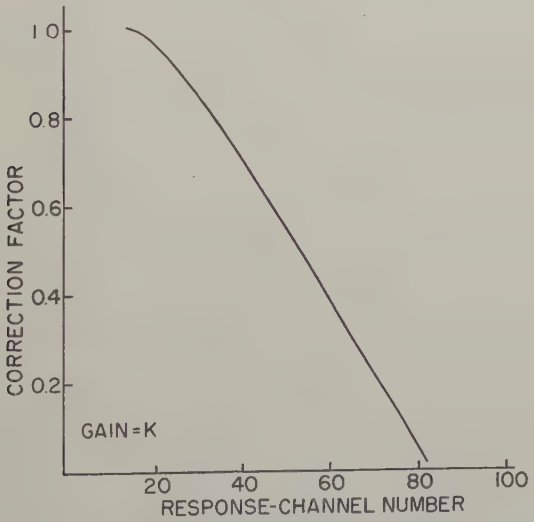
For low counter voltages with resultant low-amplitude ^{14}C pulses the new weighting factor approximates the derivative $\Delta R/\Delta S$. But as the counter voltage is increased, strong pulses tend to influence the system so that the weighting factor departs from this correction factor.

Fig. 4



Experimental correction factor versus response (channel number) for the high amplifier gain. Derivative curve is shown also.

Fig. 5



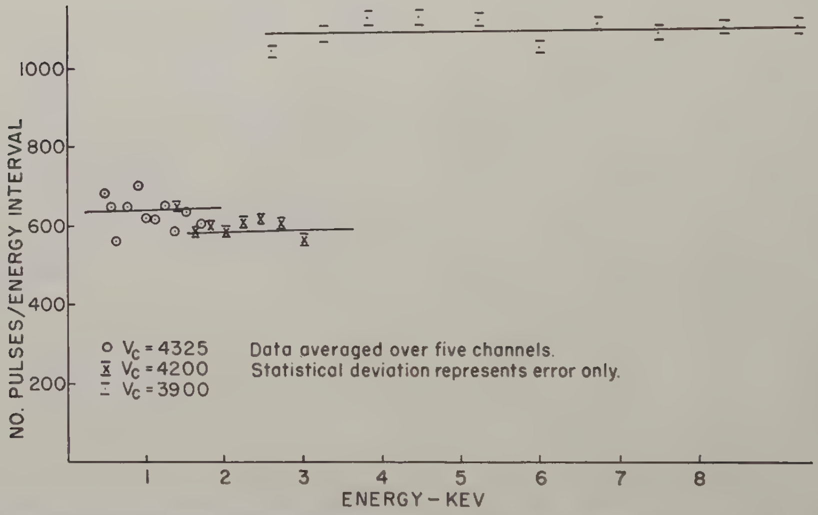
Experimental correction factor versus response (channel number) for the low amplifier gain.

Correction-factor curves for every counter voltage and counter gain used in the experiment were made with ^{14}C in the counter and again without ^{14}C. Figure 4 shows correction curves plotted against channel number for three counter voltages used with the higher amplifier setting along with the derivative curve. Figure 5 shows the curve for the lower amplification setting and corresponding counter voltage. All of the correction-factor curves taken with no ^{14}C in the counter, measuring background only, show little departure from the derivative curve. The same total disintegration rate for ^{14}C was used throughout these calibration preparations as was used in actual data runs.

§ 4. Results

Three low-energy experimental curves are shown in fig. 6. The low-energy spectrum is displayed from 0·5 to 10 kev in these curves. Each curve represents a different voltage on the main-counter anode while other parameters are held fixed thereby covering different energy regions. The different rates for these curves are due to different running times. These three sets of data are normalized and plotted in fig. 7. A theoretical intensity distribution is drawn through the data. This theoretical curve neglects the effects of screening and an approximate coulomb correction

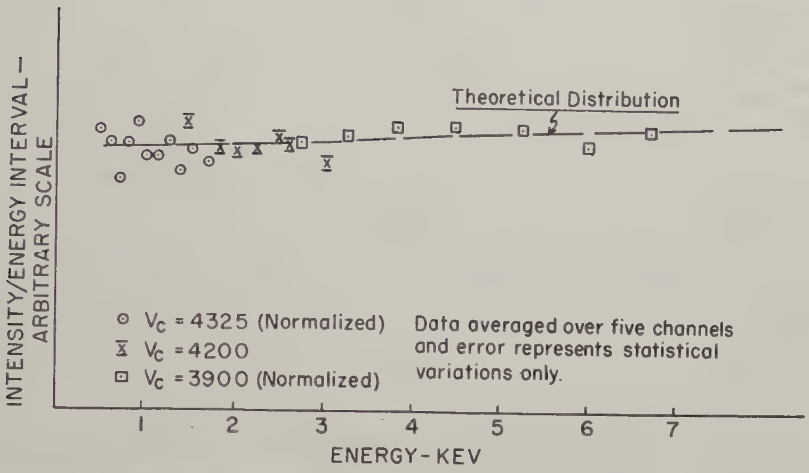
Fig. 6



Experimental intensity distributions for ^{14}C in the low-energy region. The anode voltage is varied while other parameters are held constant. The difference in rates for the curves is due to different running times. Each point represents the average of five channels in the pulse-height analyser.

factor formula is used (Kurie *et al.* 1936). These experimental data were prepared by averaging five channels in the pulse-height analyser and then subtracting the background data.

Fig. 7



Experimental intensity distribution for ^{14}C taken from the data in fig. 6 and normalized. The solid curve is a theoretical intensity distribution.

§ 5. Conclusion

Some deviations of experimental points from the allowed spectrum of ^{14}C are noted for energies greater than 7 kev. For these higher energies errors arise due to the exclusion of beta-particles of long range that cross the boundary between the main proportional counter and the guard counters. Other types of beta-ray spectrometers, however, which are accurate for this high-energy region show good agreement with theory. For energies from 0·5 kev to 7 kev the results of the present study show good agreement with theory. Thus, one may conclude that the allowed spectrum of ^{14}C with coulomb correction is in agreement with experimental data for all energies above 0·5 kev.

References

Angus, J., Cockroft, A. L., and Curran, S. C., 1949, *Phil. Mag.*, **40,** 522.
Bearden, J. A., 1933, *Rev. sci. Instrum.*, **4,** 271.
Chase, R. L., and Higinbotham, W. A., 1952, *Rev. sci. Instrum.*, **23,** 34.
Kurie, F. N. D., Richardson, J. R., and Paxton, P. C., 1936, *Phys. Rev.*, **49,** 368.
Johnstone, C. W., 1955, *The Model 250 Amplifier*, Los Alamos Scientific Laboratory Report, LA–1878, 16.
Moljk, A., Drever, R. W. P., and Curran, S. C., 1957, *Proc. roy. Soc. A*, **239,** 433.

Observations on Charged Dislocations in Ionic Crystals

By J. VENNIK, G. REMAUT and W. DEKEYSER
Laboratorium voor Kristallografie en Studie van Vaste Stoffen, Rozier, 6
Ghent, Belgium

[Received April 4, 1961]

ABSTRACT

Further results concerning charged dislocations in NaCl and LiF are described. These observations differ from those published earlier owing to the fact that the duration of the deformation pulses could be changed. It was found that in most of the specimens the dislocations were negatively charged and a 'barrier formation' is described. A second series of experiments with an Instron testing machine showed that an electrical effect could be obtained in the elastic region of the specimen.

§ 1. INTRODUCTION

IN order to explain the trapping of interstitial ions by dislocations in the silver halides, Seitz (1951) proposed that jogs on these dislocations should carry an effective charge. This charge was calculated by Seitz (1951) and found to be $\pm e/2$ for jogs in crystals having a NaCl structure.

Eshelby *et al.* (1958) then proposed that, as result of this charge on dislocations, plastic deformation of an ionic crystal would probably give an electrical effect. In three previous papers (Amelinckx *et al.* 1959, Remaut *et al.* 1960, Remaut and Vennik 1961) we described such an effect, obtained during the deformation of NaCl. From these experiments we were able to conclude that the obtained electrical effect was closely related to the movement of dislocations.

§ 2. EXPERIMENTAL

2.1. *Apparatus*

In these experiments the specimens used were rectangular slabs (3 mm × 3 mm × 6 mm) of NaCl cut from crystals grown in our laboratory, and LiF crystals bought at the Harshaw Chemical Co. The apparatus used for a first series of observations was the same as that described in a previous paper (Remaut and Vennik 1961). The specimens were mounted on a vibrator to which a series of saw-tooth current pulses was fed. The electrical signal was detected through a small electrode (0·2 mm gold wire) brought in the vicinity of that part of the surface to be investigated, and connected to a dual-trace oscilloscope. By connecting, the saw-tooth current pulse, fed to the vibrator with the second channel of the oscilloscope, it was possible to compare the obtained electrical effect with the force on the specimen.

The duration of the pulses, as well as the time between them, could now be changed within a range going from several seconds (± 5 sec) to about 10^{-3} sec. This made it possible to observe some important effects which were overlooked during the preceding observations.

A second series of experiments was performed with an Instron testing machine. A special jig was constructed for this machine so that the specimens could be periodically compressed and extended. On the Instron, one could then follow the deformation hysteresis loop of the crystal, while the corresponding electrical effect was measured on a second recorder.

2.2. *Observations*

The results for NaCl and LiF were the same, except that for the latter the electrical signal was stronger, and the effect corresponding with the elastic recoil described in a previous paper (Remaut and Vennik 1961) was bigger.

An electrical signal was always obtained, independently of the duration of the pulses used. We have not been able to find, at room temperature, a critical rate of deformation with which no electrical effect would be observed. If the specimen was deformed by two pulses of the same amplitude but of different duration, then the signal corresponding to the pulse with the longest duration was the strongest. This effect is illustrated in figs. 1 and 2†. The pulse in fig. 2 has the same amplitude as the pulse in fig. 1, but its duration in ten times smaller. As one can see, its corresponding signal is also smaller than that corresponding to the pulse in fig. 1. In fig. 3 the relationship between the duration of the pulses and the intensity of the signal is shown for a specimen. The signal was independent of the dead time between the deformation pulses.

As was stated in a previous paper (Remaut and Vennik 1961) if the specimen is deformed by a series of pulses of the same polarity and amplitude, then, after the first pulse, the signal does not changed anymore. i.e. a d.c. voltage is obtained which is approximately equal to the maximum value of the signal corresponding to the first pulse. But if the specimen is deformed by a series of pulses of constant polarity and increasing amplitude, then one observes that the d.c. voltage corresponding to a particular pulse starts increasing when the maximum stress of this pulse is reached by the following one.

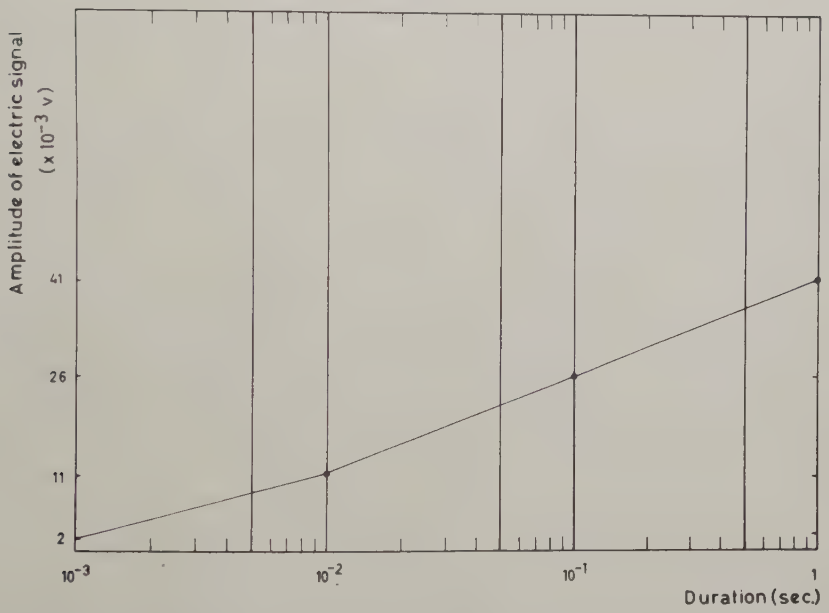
Periodic deformation of the specimen with pulses of alternating polarity and constant amplitude gives a periodic signal of which the amplitude decreased with time. After 10 to 20 pulses an equilibrium amplitude was obtained. Its value depended on the amplitude of deformation and on the history of the specimen. The shape of the signal changed also. From being originally concave upwards, the end part became concave downwards with respect to the time axis. If, after this equilibrium signal was reached, the amplitude of deformation was suddenly increased then a strong effect on the observed signal was seen. This is shown in figs. 4 and 5.

---

† Figures 1, 2, 4 and 5 are shown as plates.

In fig. 4 one has the electrical signal just before the amplitude of deformation was increased. In fig. 5 the electrical signal is shown just after the amplitude of deformation was increased. The signal is now very irregular. This effect must be the result of a change in the speed of the dislocations or of the charge on the dislocations or both. In principle, the reduction in speed should be detectable with a sensitive tensile machine.

Fig. 3



Relationship between the duration of the pulses and the intensity of the signal.

Everything seems to suggest that, as a result of the alternating pulses, some kind of barrier for the dislocations within the crystals is formed. By augmenting the amplitude of deformation, the dislocations are made to break through this barrier and this results in an irregular shaped signal.

From the photos taken, one can see that this break-through takes place when the maximum stress of the preceding pulse is reached. One can also notice a decrease in the slope of the signal just before the break-through, and a sharp increase in the slope just after the break-through.

Although about 10 to 20 pulses are needed to form this barrier completely, it is already noticeable after 3 to 4 pulses. The influence of this barrier on the signal could be seen for about 6 pulses after the break-through. By then, this obstacle seems to have been cleaned up, and the signal took on its equilibrium form again. While this barrier was being cleaned up, a second barrier was formed further up in the slip plane. By increasing for the second time the deformation amplitude, one got a new break-through.

It was found that such a barrier still existed after two days at room temperature. The intensity of the barrier was also influenced by time between the pulses. The longer the time, the stronger the barrier for a given number of cycles.

To see if this barrier formation was not a surface effect, a surface layer of a specimen in which a barrier had been formed was taken away with water and then dried with alcohol and ether. When the amplitude of deformation was increased the barrier was still observed. Pulses with the same polarity did not form such a barrier, nor did they clean up an existing one.

In nearly all of our experiments, and contrary to what was stated in a preceding paper (Remaut *et al.* 1960), the dislocations were found to be negative. The observations with the Instron showed that a small electrical effect was already obtained in the elastic region of the specimen. In the plastic region the signal increases suddenly to high values. We therefore kept the deformation for most of our experiments around the yield point. In this way the signal could be kept under control, and the life of the specimen used was nearly unlimited.

During a few experiments with the Instron testing machine, a changing of the sign of the electrical effect was observed. As this always happened when the amplitude of deformation was high, we assumed that this resulted from slip along glide systems different from those already active.

§ 3. DISCUSSIONS

From these experiments one may conclude that:

(*a*) An electrical effect is already obtained in the elastic region of a specimen.

(*b*) A ‘ barrier formation ’ or ‘ memory effect ’ of the specimen results when the crystal is submitted to a series of pulses of constant amplitude and alternating polarity.

(*c*) Moving dislocations behave as negatively charged in NaCl and LiF. As to the sign of the dislocations the difficulty is to explain the negative charge found on them, contrary to what was predicted by Eshelby *et al.* (1958). In this respect it seems logical to consider first of all the role of positive-ion vacancies present in our specimens. If they play a prominent role, experiments with doped crystals should have a marked effect on the intensity or the sign of the signal.

To test this assumption a series of observations was performed on specimens doped with Cd^{++} ions. These experiments showed that the concentration of the positive-ion vacancies had no effect on the sign of the signal and on its magnitude.

One could object that our pure crystals contained already a number of impurities sufficient to make any further doping ineffective. However, other tests on these crystals, such as ionic conductivity, absorption spectra after irradiation, etc., showed a marked influence of impurity addition. This indicates that the uptake of positive-ion vacancies by moving dislocations must not be considered as an explanation for the origin and sign of this effect.

On the other hand, the observed barrier formation seems to indicate that vacancies cannot be excluded. Therefore it is possible that the negative sign of the dislocations results from the boiling off of an excess negative-ion vacancies during the movement of the dislocations, and this contrary to what one would expect from the calculations by Mott and Littleton (1938) concerning the formation-energy of negative-ion vacancies.

We have at the moment no definite evidence for such a mechanism and experiments are in progress to ascertain this hypothesis. The effect of temperature will in particular be examined.

ACKNOWLEDGMENT

This work is part of a research scheme supported by the Institut pour l'Encouragement de la Recherche Scientifique dans l'Industrie et l'Agriculture (C.E.S.).

REFERENCES

AMELINCKX, S., VENNIK, J., and REMAUT, G., 1959, *J. Phys. Chem. Solids*, **11,** 170.

ESHELBY, S. D., NEWEY, C. W. A., PRATT, P. L., and LIDIARD, A. B., 1958, *Phil. Mag.*, **3,** 76.

MOTT, N. F., and LITTLETON, N. J., 1938, *Trans. Faraday Soc.*, **34,** 485.

REMAUT, G., VENNIK, J., and AMELINCKX, S., 1960, *J. Phys. Chem. Solids*, **16,** 158.

REMAUT, G., and VENNIK, J., 1961, *Phil. Mag.*, **6,** 1.

SEITZ, F., 1951, *Rev. mod. Phys.*, **23,** 328.

Ion Damage to Metal Films inside an Electron Microscope

By D. W. Pashley and A. E. B. Presland†
T.I. Research Laboratories, Hinxton Hall, Cambridge

[Received January 30, 1961]

Abstract

The origin of small spot and loop features observed in electron micrographs of thin evaporated single-crystal gold films has been investigated. It is found that the features are introduced into the specimens whilst they are being examined in the electron microscope. Conclusive evidence is given to show that they arise from bombardment by energetic negative ions emitted from the tungsten filament of the electron gun, although the ions have not been identified. The rate of ion damage can be considerably increased by coating the tungsten filament with a standard oxide emitter.

The annealing characteristics of the damage have been studied. Above about 300–350°c the majority of the spot features anneal out, and ion damage above about 350°c consists of the formation of small tetrahedra of stacking fault.

The ion damage is assumed to be very similar in character to that produced by primary knock-ons resulting from irradiation by other particles (e.g. neutrons).

§ 1. Introduction

During the past five years the electron microscope has been used in this laboratory to examine in detail the growth and microstructure of thin single-crystal films prepared by evaporation (for reviews see Bassett and Pashley 1958–59, Bassett *et al.* 1959). One of the microstructural features which has been observed in the films, particularly in gold, is a small dot-like structure (Pashley 1959, Pashley and Presland 1961), the density of which appeared to vary considerably from specimen to specimen. The dots often have the appearance of small loops (30 Å–100 Å in diameter), similar to the dislocation loops first observed in quenched aluminium by Hirsch *et al.* (1958). The general appearance of the structure bears a remarkable resemblance to that observed in neutron-irradiated copper by Silcox and Hirsch (1959 b). These authors use the term ‘ spots ’ to describe the appearance of the structure; we shall adopt the same terminology here, without necessarily implying that such features are always of the same nature. Observation of spots in single crystal films prepared by evaporation has also been reported by Matthews (1959) and Phillips (1960).

The purpose of this paper is to describe an investigation into the origin of the spots observed in the evaporated films. Included in the work is a study of their annealing characteristics. It is shown that the spots are introduced into the specimen whilst it is under observation in the electron microscope, and that they arise from bombardment by negative ions produced in the electron gun.

---

† Now at C. A. Parsons & Co. Ltd., Nuclear Research Centre, Newcastle upon Tyne.

§ 2. EXPERIMENTAL TECHNIQUES

The specimens used in the investigation were single-crystal films of gold (100 Å–500 Å in thickness) prepared by evaporation onto silver/mica ((111) orientation) and silver/rockslat ((100) orientation) substrates (Pashley 1959). The films were well oriented and very coherent, approximating well to parallel-sided slabs. They have in all cases been examined in transmission in a Siemens Elmiskop I operating at 80 or 100 kv.

In addition to the normal observations at room temperature the specimens have also been examined on a hot-stage previously decsribed (Pashley and Presland 1958–59), and on a stage of the type first constructed by Heide (1960). This latter stage allows the immediate neighbourhood of the specimen to be surrounded by a gas at a pressure considerably greater than that of the main vacuum, and can be used to prevent carbon contamination on the specimen. A description of this apparatus will be given elsewhere.

§ 3. THE FORMATION OF THE SPOT STRUCTURE INSIDE THE ELECTRON MICROSCOPE

Systematic experiments have shown that the spots appear whilst the specimen is under observation with the electron beam. The rate of formation of the spots varies considerably; sometimes it is so fast that it is extremely difficult to adjust the controls to obtain a good image of the specimen before quite a high density is present. The number of distinguishable individual features produced varies between the approximate limits 10^7–10^9/cm^2/sec, the lower value representing the more normal rate. Figure 1 (*a*), (*b*)† shows different areas of the same specimen after periods of ~ 100 sec and ~ 1000 sec of observation, and corresponds to an approximate rate of production of 5×10^8/cm^2/sec. Figure 2 is an example of the spots in which the loop character is well illustrated (e.g. at A). The detailed appearance of the spots depends upon the diffracting conditions, but no systematic study of this has been made as yet. It seemed likely that the spot structure is due to the aggregation of point defects, and two general mechanisms are possible. Firstly, the point defects could be introduced into the specimens during preparation, and their aggregation brought about during examination. Secondly, both the creation and aggregation could take place during the examination. Since the heating effect of the electron beam seemed to be one possible case of the first mechanism, the specimens were deliberately annealed inside the electron microscope, in order to test this hypothesis.

3.1. *Annealing Experiments*

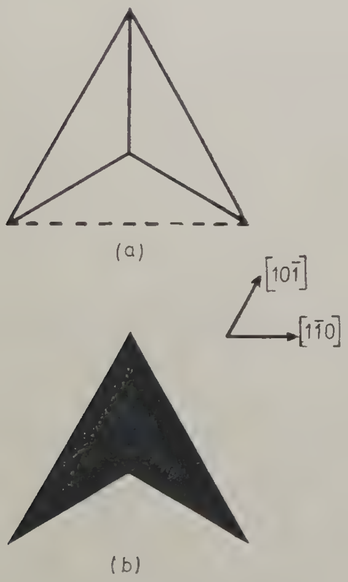
Two kinds of annealing experiment have been carried out. In the first the specimen was examined (i.e. the electron beam was switched on) only while it was maintained at various temperatures up to 400°c. In the second, the spots were produced in the specimen at a certain temperature and the effect of raising the temperature was studied.

---

† Figures 1–5 and 7–9 are shown as plates.

It was found that annealing alone did not form spots in the gold films. Although spots were formed in the specimens at all temperatures up to about 300°c, they only appeared as a result of having the electron beam switched on. An example of the structure produced at elevated temperatures is shown in fig. 3. Quite large dislocation loops were sometimes formed (e.g. at B). Above about 300–350°c no spots were formed and any previously produced structure was annealed out (fig. 4). However, at temperatures of 350–400°c new features such as shown in fig. 5 appeared.

Fig. 6



The contrast due to a tetrahedron of stacking fault, projected on the (111) plane. (*a*) the stair-rod dislocations in contrast (full line), and out of contrast (broken line), for the $(2\bar{2}0)$ reflection operating, (*b*) the stacking faults in contrast under the same conditions.

These are interpreted as tetrahedra of stacking fault, as first observed in quenched gold by Silcox and Hirsch (1959 a). Figure 6 (*a*) shows a tetrahedron of stacking fault projected on the (111) plane. When a single strong (220) reflection is operating, the stair-rod dislocations shown by full lines in fig. 6 (*a*), and the stacking faults forming the two inclined faces indicated in fig. 6 (*b*), are in good contrast (for details, see Appendix). It follows that the arrow head nature of the contrast (see fig. 5) will be observed whether it is due to the stair-rod dislocations or to the stacking faults. A more detailed treatment is required to take into account the combined effect of the two. However, this general interpretation has been confirmed by means of dark field experiments which show that the direction of the arrow-head is related to the reflection operating as indicated in fig. 6, for the case of the $(2\bar{2}0)$ reflection. Further, there are two distinct orientations of the tetrahedra (double-positioning) as shown in fig. 5. This is expected from purely geometrical considerations.

These results lead to the conclusion that aggregates of point defects are formed duirng the examination of the gold films in the electron microscope. It was further established that the point defects are actually introduced during observation, and not during the preparation of the specimens, because when any of the specimens annealed at temperatures above 300°c were cooled to room temperature (with electron beam switched off), the spots reappeared when the electron beam was switched on. Since, when the spots are removed by annealing above 300°c, the emitted point defects should escape at the film surface, the spots which reappeared on subsequent cooling must have arisen from freshly created point defects.

This conclusion suggests that the formation of the spots should be largely independent of the method of preparation of the gold film. It was in fact found that spots are also produced in a similar manner in a specimen prepared by the standard technique of electro-thinning of a gold foil originally about 0·001 in. in thickness (see fig. 7). The detection of the spots was more conclusive in the thinner regions (< 1000 Å) of the specimen.

§ 4. THE POSSIBILITY OF DIRECT ELECTRON EFFECTS

Electrons of energy 80 kev should not create Frenkel defects in gold, as the threshold energy is of the order of 1 Mev. It seemed possible that special conditions at the film surface might allow point defects to be introduced into the interior of the specimen. For example, it is well known that a specimen rapidly becomes coated with a carbon contamination film whilst it is being examined in an electron microscope. It so happens that the temperature dependence of the rate of carbon contamination correlated approximately with the temperature dependence of the rate of production of the spots. Experiments were therefore carried out to check whether this correlation is significant. The method described by Heide (see § 2) was used to allow the gold films to be examined at room temperature in the complete absence of any carbon contamination. It was found that the spots still formed.

A direct check that the electron beam is not responsible, either directly or indirectly, for the formation of the spots was made in the following way. A pair of deflecting coils was mounted just above the specimen, and the d.c. current through them was adjusted so that the electron beam (focused to a $\sim 2\,\mu$ spot) was deflected right off the specimen. A fresh specimen was inserted into the microscope and the electron beam was switched on only after the deflecting coil current had been switched on. After about one hour, the deflecting coil current was switched off so that a good image of the specimen was obtained within a few seconds. It was found that the spots had formed with a high density even with no electron bombardment of the specimen. This is consistent with the general experience that the spots are not confined to the area of the specimen where the electron intensity is greatest.

A possible interpretation of these results seemed to be that energetic negative ions produced in the electron gun are responsible for the formation of the spots. The relatively heavy ions would be much less deflected by either the condenser lens fields or the above deflecting field. For example, the focal lengths of the two condenser lenses, under normal illuminating conditions, are ~ 50 and ~ 1000 metres respectively for 80 kev singly charged oxygen ions. Furthermore, 80 kev ions have ample energy for producing considerable numbers of point defects. This possibility is now considered in detail.

§ 5. TESTS OF THE ION HYPOTHESIS

If negative ions are produced in the electron gun, particularly at or in the immediate vicinity of the tip of the filament, they will be accelerated by the applied gun field and have a significantly high probability of reaching the specimen. However, we know of no previous evidence that significant numbers of negative ions are produced by guns having non-coated filaments.

The penetration of ions into metals is very much less than that of the electrons. According to the formulae used by Gillam (1959) and derived by Nielson (1956), the penetration of 100 kev argon ions (for example) into gold is about 350 Å, and so the majority of such ions would be filtered out by a film of this thickness. The filtering action of the gold films has therefore been tested. Two gold films of 200–300 Å in thickness have been mounted one above the other, about 1 mm apart, so that both could be brought into focus in turn. Preliminary tests showed that whereas the image of the lower film was adequate to allow the spots to be detected, the image of the upper film, viewed through the lower film, was not sufficiently good for the purpose. A hole was therefore made in the lower film, so that part of the upper film could be examined by electrons which had not passed through the lower film. It was found that the normal high spot density was formed in the upper film (see fig. 8 (*a*)), but even after the beam had been switched on for $1\frac{1}{2}$ hours, a negligible number appeared in the lower film (see fig. 8 (*b*)).

This result, together with the deflection experiment described above (§ 4), shows that the spots are formed by radiation which passes down the electron microscope column, is heavily absorbed in the metal film and is much less deflected by a magnetic field than the 100 kev electrons. This is taken as extremely strong evidence that the damage is caused by negative ions produced in the electron gun.

It is well known that oxide-coated emitters (e.g. in cathode-ray tubes) give rise to negative ion beams of relatively high intensity. Tests have therefore been carried out with normal tungsten filaments for the Siemens

Elmiskop I coated with a standard emitter containing barium and strontium salts. A good bright image could be obtained with these filaments when the filament voltage was very much lower than normal. It was found that the spots appeared at an extremely enhanced rate (at least $10^{11}/\mathrm{cm}^2/\mathrm{sec}$) compared with normal. Figure 9 (*a*), (*b*), (*c*) shows examples of the structure produced after periods of irradiation of no more than about one minute. The instantaneous appearance of the individual spots could be readily observed on the fluorescent screen of the microscope, many appearing simultaneously within the field of the view when the screen magnification was 40 000 × . The form of the spots, except at heavy doses, is indistinguishable from that produced at the lower rate under normal operating conditions (cf. figs. 1 (*b*) and 9 (*a*)). Tetrahedra of stacking fault could be detected on some of the images (e.g. at C, fig. 9 (*b*)). Figure 9 (*b*) and (*c*) shows effects, such as at D, associated with heavy doses. They are believed to be due to the coalescence of spots formed in close proximity to each other.

From this experiment it is concluded that the normally produced spots result from bombardment by negative ions produced in the electron gun.

§ 6. Discussion

The results presented in this paper show quite conclusively that damage to gold films occurs in specimens whilst they are under observation in the electron microscope. This damage is almost certainly due to bombardment by negative ions produced in the electron gun, although the mechanism by which the ions are produced is not known. The tungsten filament of an electron gun is continuously thinning during operation, and this is believed to be due in part to the formation of a volatile oxide at its surface. Possibly oxygen ions could be produced during the evaporation of this oxide. Alternatively, the ions might originate from impurities present in the tungsten wire, since we have shown that by adding other substances to the surface of the wire it is possible to considerably enhance the rate of ion production (see § 5). The production of sufficient numbers of negative ions in the residual gas of the electron gun seems much less plausible, since only a small proportion of these would be accelerated in the right direction to allow them to reach the specimen.

The observed considerable variation in the rate of production of spots (see § 3), which is presumably due to variation in ion current density at the specimen level, can now be explained, in two general ways. Either the actual ion source is variable, or big variations occur in the efficiency with which the ions reach the specimen. No serious attempt has as yet been made to distinguish between these two possibilities, although it is clearly of some importance in relation to the behaviour of other microscopes with respect to specimen ion damage. Possible causes for the variation in the tungsten filament as an ion source are (*a*) variation of impurity content, and (*b*) variation in oxidation rate due to changes in the residual vacuum in the gun. Possible causes for the variation in the efficiency with which ions reach the specimen are (*a*) variation in gas pressure causing changes in the scattering of the ions by the gas molecules, and (*b*) variation in the alignment of the illuminating system of the microscope. This latter effect should not be too serious, because there are only two effective restricting apertures in the Siemens Elmiskop I, a fixed one of about 750 μ diameter, and an adjustable one which is normally chosen to be 200–600 μ in diameter. At worst, one would expect the area of the ion damage (at least 300–900 Å diameter, depending upon the size of the adjustable condenser aperature) to be not at the centre of the specimen, but shifted to one side. It seems unlikely that misalignment would lead to complete protection of the specimen from ions.

Although the main observations have been made on gold films, similar ion damage has been observed in Ag, Cu, Pd, Pt and Rh. The nature of the ion damage is such that it is almost certainly due to the aggregation of point defects produced by the ions. The individual spots are seen to appear suddenly, and usually singly, suggesting that each feature results from the impact by one ion. However, particularly at elevated temperatures, some features (e.g. B in fig. 3) are shown to grow in size gradually suggesting that point defects produced by a number of different ion impacts eventually contribute to one feature. It is not yet clear which of these two processes is the more prominent.

Since irradiation of solids by other particles (e.g. neutrons) gives rise to secondary atoms or ions (primary knock-ons) inside the solid, and since much of the final damage is produced by these secondary particles, the damage by direct ion bombardment represents a fundamental component of more general radiation damage by heavy particles. This explains the similarity between the damage observed in pile irradiated copper (Silcox and Hirsch 1959 b), and the ion damage produced in the gold films. Most theories of radiation damage predict the formation of point defects by the primary knock-ons and the aggregation of some of these defects to produce dislocation loops (e.g. Seitz 1949, Brinkman 1956).

The direct observation of ion damage in the electron microscope thus represents a powerful means of studying radiation damage, particularly as the temperature of the specimen can be independently controlled within close limits. Also damage which rapidly anneals out at room temperature can be studied in this way, whereas it is much more difficult by other techniques.

Since the main purpose of this paper is to consider the ion damage in terms of an ‘ artefact ’ in transmission electron microscopy, a more detailed consideration of the nature of the damage, and its relation to general radiation damage in crystalline solids, will be given elsewhere.

ACKNOWLEDGMENTS

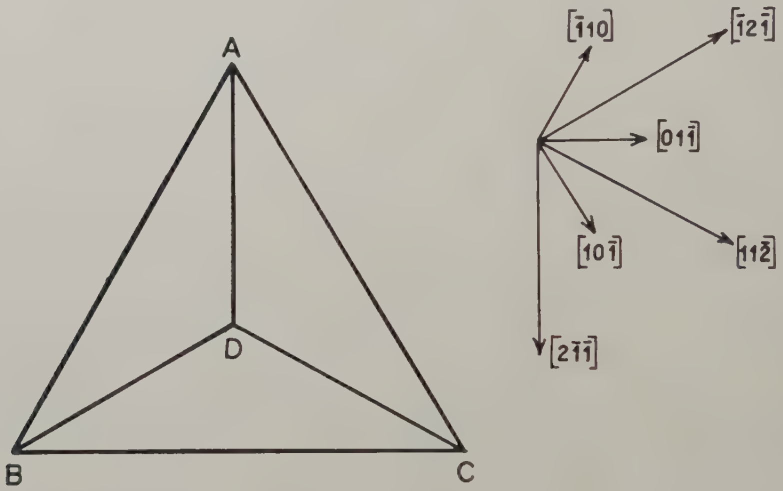
We are grateful to Dr. J. W. Menter for stimulating discussions. This paper is published by permission of the Chairman of Tube Investments Limited.

APPENDIX

The Contrast due to Tetrahedra of Stacking Fault

A tetrahedron of stacking fault is bounded by four stacking faults on its faces, and six stair-rod dislocations alongs its edges. As a first approximation we consider the contrast arising from the stair-rod dislocations and the stacking faults separately. According to the contrast mechanism discussed by Hirsch *et al.* (1960), the dislocations are in contrast provided $\mathbf{g}.\mathbf{b}\neq 0$, where $\mathbf{g}$ is the reciprocal lattice vector corresponding to the operating reflecting plane, and $\mathbf{b}$ is the Burgers vector of the dislocation. Table 1 gives the values of $\mathbf{g}.\mathbf{b}$ for the six stair-rod dislocations, the indices being as indicated by reference to fig. 10.

Fig. 10



A tetrahedron of stacking fault projected on to the (111) plane.

The values have been determined for the two principal kinds of reflections which occur with a film in (111) orientation. In both cases, all but one of the stair-rods is in contrast; the result for the $(2\bar{2}0)$ reflection operating is indicated in fig. 6 (*a*).

The contrast due to stacking faults inclined to the plane of the film has been considered in detail by Whelan and Hirsch (1957). The characteristic fringe contrast does not appear in the present case because the extinction distance (~280 Å for a (220) reflection) is greater than the size of the observed tetrahedra (about 100 Å long an edge). Instead, the entire projected area of the fault appears in near uniform contrast with its surroundings. The contrast will occur provided $\mathbf{g}.\mathbf{R}$ is not an integer, where $\mathbf{R}$ is the displacement vector corresponding to the fault. $2\pi\mathbf{g}.\mathbf{R}$

represents the phase shift of the diffracted beam on one side of the fault relative to the other, resulting from the presence of the fault. Values of this quantity are given in table 2. For the ($2\bar{2}0$) reflection operating, only two of the four faults are in contrast, as indicated in fig. 6 (*b*).

Table 1. **g** . **b** values for the stair-rod dislocations

Stair-rod	**b**	**g** . **b**	
		g$=[2\bar{2}0]$	**g**$=[4\bar{2}\bar{2}]$
AB	$\frac{1}{6}$ [110]	0	$+\frac{1}{3}$
BC	$\frac{1}{6}$ [011]	$-\frac{1}{3}$	$-\frac{2}{3}$
CA	$\frac{1}{6}$ [101]	$+\frac{1}{3}$	$+\frac{1}{3}$
AD	$\frac{1}{6}$ $[01\bar{1}]$	$-\frac{1}{3}$	0
BD	$\frac{1}{6}$ $[\bar{1}01]$	$-\frac{1}{3}$	-1
CD	$\frac{1}{6}$ $[1\bar{1}0]$	$+\frac{2}{3}$	$+1$

Table 2. 2π**g** . **R** values for the stacking faults

Face (see fig. 10)	Indices of face	**R**	2π**g** . **R**	
			g$=[2\bar{2}0]$	**g**$=[4\bar{2}\bar{2}]$
ABC	(111)	$\frac{1}{3}$ [111]	0	0
ABD	$(11\bar{1})$	$\frac{1}{3}$ [112]	0	$-\pi/3$
ACD	$(1\bar{1}1)$	$\frac{1}{6}$ [121]	$-\pi/3$	$-\pi/3$
BCD	$(\bar{1}11)$	$\frac{1}{6}$ [211]	$+\pi/3$	$+2\pi/3$

This kind of treatment of the contrast is clearly not sufficient to allow all details of the images to be explained, since the contrast due to the stair-rod dislocations and the stacking faults cannot be treated separately, particularly with such small tetrahedra. However, it is sufficient to allow reasonable identification of the stacking-fault tetrahedra, if the contrast is correlated with the reflection operating.

REFERENCES

BASSETT, G. A., and PASHLEY, D. W., 1958–59, *J. Inst. Met.*, **87,** 449.

BASSETT, G. A., MENTER, J. W., and PASHLEY, D. W., 1959, *Structure and Properties of thin Films* (New York: John Wiley and Sons Inc.), p. 11.

BRINKMAN, J. A., 1956, *Amer. J. Phys.*, **24,** 246.

GILLAM, E., 1959, *J. Phys. Chem. Solids*, **11,** 55.

HEIDE, H. G., 1960, *Fourth International Congress on Electronmicroscopy, Berlin* (Berlin, Gottingen, Heidelberg: Springer), p. 82.

HIRSCH, P. B., HOWIE, A., and WHELAN, M. J., 1960, *Phil. Trans.*, **252,** 499.

HIRSCH, P. B., SILCOX, J., SMALLMAN, E. R., and WESTMACOTT, K. H., 1958, *Phil. Mag.*, **3,** 897.

MATTHEWS, J. W., 1959, *Phil. Mag.*, **4,** 1017.

NIELSON, K. O., 1956, *Electromagnetically Enriched Isotopes and Mass Spectrometry* (London: Butterworths), pp. 68–81.

PASHLEY, D. W., 1959, *Phil. Mag.*, **4,** 324.

PASHLEY, D. W., and PRESLAND, A. E. B., 1958–59, *J. Inst. Met.*, **87,** 419; 1961, *The Proceedings of the European Regional Conference on Electron Microscopy, Delft*, 1960 (De Nederlandse Vereniging Voor Electronenmicroscopie), Delft, 1961, p. 391.

PHILLIPS, V. A., 1960, *Phil. Mag.*, **5,** 571.

SEITZ, F., 1949, *Disc. Faraday Soc.*, **5,** 271.

SILCOX, J., and HIRSCH, P. B., 1959 a, *Phil. Mag.*, **4,** 72; 1959 b, *Ibid.*, **4,** 1356.

WHELAN, M. J., and HIRSCH, P. B., 1957, *Phil. Mag.*, **2,** 1121.

A Theory of the Electrical Properties of Liquid Metals
I: The Monovalent Metals

By J. M. Ziman
Cavendish Laboratory, University of Cambridge

[Received May 23, 1961]

> *And, sure, the reverent eye must see*
> *A purpose in liquidity.*
> Rupert Brooke.

Abstract

The mean free path of an electron in a liquid metal depends on scattering by the ions. The cross section of each ion depend on its ' pseudo-potential ', which can be estimated from the band gap in the solid. If the ions scattered independently, the mean free path would be much too short. But there is strong correlation between the positions of the ions in the liquid, giving rise to coherence between waves diffracted by adjacent ions. Using simple perturbation theory, one can show that the resistance should depend on the Fourier transform, $a(K)$, of the radial distribution function, and should be smaller than for independent scatterers.

The density fluctuations of the liquid also scatter electrons, through imperfect screening by the ' other electrons ' of the coulomb fields of the ions. This ' plasma resistance ' is the main effect in liquid Na, where the band gap is small. When this term has been subtracted, the resistivities of the other monovalent metals come out according to the theory, being approximately proportional to the squares of their pseudo-potentials.

The change of resistance on melting is shown to be due to the change in the radial distribution function. Different ratios of ρ_L/ρ_S can be explained through differences in the scattering cross sections of ions of different metals —especially the differences in the ion core ' pseudo-potentials '. The temperature variation of the resistivity follows from the temperature dependence of $a(K)$. The thermoelectric power in the solid, and in the liquid, and its change on melting, all follow qualitatively from the form of $a(K)$ and of the scattering cross section for each type of ion.

---

§ 1. The Pseudo-potential

A fruitful concept that is now fashionable in the theory of electronic structure of metals is the notion of the ' pseudo-potential ' (Phillips and Kleinman 1959). This is a simple idea in principle, although not easy to justify precisely (Cohen and Heine 1961). We assert that the conduction electrons behave as if they were ' free ', except for a modest field attached to each ion. This field represents a balance between an electrostatic attraction and the ' Pauli repulsion ' exerted by the bound states in the core, to which the conduction electron must be orthogonal. The deep singularity in the electrostatic term inside the ion core is very nearly cancelled out by the large kinetic energy associated with the rapid oscillations of the wave function in this region, so that the net effect is the same as a weak pseudo-potential acting on a 'pseudo wave-function', which is the same as the real wave function without the rapid oscillations in the core. Although some caution is needed in accepting too literally a single function $\mathscr{U}_a(\mathbf{r}-\boldsymbol{l})$ to represent the pseudo-potential for all electronic states (cf. Ham 1960) it does seem to offer a fair picture of the energies and wave functions in a single band, or in the restricted range of states associated with the electrons on the Fermi surface in the monovalent metals (Ziman 1961, Collins 1961, Collins and Ziman 1961).

Let us recall the next steps in the analysis for a perfect crystal. The total potential field acting on a conduction electron at $\mathbf{r}$ would be the sum of contributions from the ions at their lattice sites $\boldsymbol{l}$:

$$\mathscr{U}(\mathbf{r})=\textstyle\sum_l \mathscr{U}_{\mathrm{a}}(\mathbf{r}-\boldsymbol{l}). \qquad \qquad (1)$$

We treat this as a perturbation on our plane waves states—assumed to be of the form $V^{-1/2}\exp(i\mathbf{k}\,.\,\mathbf{r})$ (which now replace the OPW's in this formalism). We need the matrix elements between the unperturbed states:

$$\begin{aligned}\mathscr{U}_{\mathbf{k},\,\mathbf{k}'}&=\frac{1}{V}\int \exp(-i\mathbf{k}\,.\,\mathbf{r})\mathscr{U}(\mathbf{r})\exp(i\mathbf{k}'\,.\,\mathbf{r})\,d\mathbf{r}\\ &=\frac{1}{V}\textstyle\sum_l \displaystyle\int \exp[i(\mathbf{k}'-\mathbf{k})\,.\,\mathbf{r}]\mathscr{U}_{\mathrm{a}}(\mathbf{r}-\boldsymbol{l})\,d\mathbf{r}\\ &=\frac{1}{V}\textstyle\sum_l \exp[i(\mathbf{k}'-\mathbf{k})\,.\,\boldsymbol{l}] \displaystyle\int \exp[i(\mathbf{k}'-\mathbf{k})\,.\,\mathbf{r}]\mathscr{U}_a(\mathbf{r})\,d\mathbf{r}\\ &=\mathscr{U}_{\mathbf{g}}\delta_{\mathbf{k}'-\mathbf{k},\,\mathbf{g}} \qquad \qquad (2)\end{aligned}$$

where $\mathbf{g}$ is any vector of the reciprocal lattice, and $\mathscr{U}_{\mathbf{g}}$ is the corresponding Fourier component of $\mathscr{U}_a(\mathbf{r})$, i.e.

$$\mathscr{U}_{\mathbf{K}}\equiv\frac{N}{V}\int \exp(i\mathbf{K}\,.\,\mathbf{r})\mathscr{U}_a(\mathbf{r})\,d\mathbf{r}. \qquad \qquad (3)$$

Now we can construct wave functions and energies by perturbation theory. Thus, we expand the wave function in the form

$$\psi_{\mathbf{k}}=\textstyle\sum_{\mathbf{g}}\alpha_{\mathbf{g}}\exp[i(\mathbf{k}+\mathbf{g})\,.\,\mathbf{r}] \qquad \qquad (4)$$

where

$$\alpha_{\mathbf{g}}\approx\sum_{\mathbf{g}}\frac{\mathscr{U}_{\mathbf{g}}}{\mathscr{E}_{\mathbf{k}}{}^0-\mathscr{E}_{\mathbf{k}+\mathbf{g}}{}^0} \qquad \qquad (5)$$

and

$$\mathscr{E}_{\mathbf{k}}\approx\mathscr{E}_{\mathbf{k}}{}^0+\sum_{\mathbf{g}}\frac{|\mathscr{U}_{\mathbf{g}}|^2}{\mathscr{E}_{\mathbf{k}}{}^0-\mathscr{E}_{\mathbf{k}+\mathbf{g}}{}^0}+\ldots. \qquad \qquad (6)$$

We take the unperturbed energies, of course, to be just $\mathscr{E}_{\mathbf{k}}{}^0=\frac{1}{2}k^2$ and assume that $\mathscr{U}_0$ is zero, or has already been added to all the unperturbed energies.

This expansion will work only when there is no degeneracy. If, for any value of $\mathbf{g}$, it should happen that $\mathscr{E}_{\mathbf{k}}{}^0=\mathscr{E}_{\mathbf{k}+\mathbf{g}}{}^0$—that is to say, if $\mathbf{k}$ lies on the zone boundary that bisects the reciprocal lattice vector **g**—then we must evaluate the secular determinant for these states and choose the appropriate roots for the energy. It is well known that states on a zone *face* (not at edges or corners, where the situation is more complex) are then split into two levels by an amount

$$\mathscr{E}_{\text{gap}} = 2|\mathscr{U}_{\mathbf{g}}|. \qquad (7)$$

In the monovalent metals there is a gap at the nearest face of the zone—the (111) face or the (110) face, according to the crystal structure—which determines the shape of the Fermi surface, etc. From the transport properties it is possible to estimate the size of the gap (Cohen and Heine 1958, Ziman 1961, Collins and Ziman 1961), and there have recently been some direct band-structure calculations (Ham 1960, Segall, private communication) which confirm these estimates. The data are not very reliable, but show quite clearly that $\mathscr{E}_{\text{gap}}$ is nearly zero for Na, and perhaps also K, increases in magnitude as we go up to Li, or down to Rb and Cs, and is larger still in the noble metals, Cu, Ag, Au, where the Fermi surface is drawn into actual contact with the zone boundary.

This is our primary information; we have some idea of the magnitude of one Fourier component of the pseudo-potential in each of the monovalent metals. This is not the whole function $\mathscr{U}_{\mathbf{K}}$, but at least we may reasonably assume that its profile is much the same in all cases, and only its magnitude varies from one metal to another. As for this profile, it can only be said that the calculation of Phillips and Kleinman (1959) indicated fairly large Fourier components at (111) and (200) in their calculation for Si, with smaller values thereafter—indeed, as shown by Cohen and Heine (1961), tending exactly to cancel at large wave numbers.

§ 2. Gas Approximation

The prime characteristic of a liquid is that it is disordered. Let us take this to its logical limit, and suppose that it is thoroughly disordered—that is, that there is no correlation between the positions of the ions. In effect, we treat the ions as a gas of point particles, which may approach each other as close as you please. The ions acts as independent scattering centres which limit the mean free path of the electrons. It is easy to prove (e.g. Ziman 1960, § § 7.1, 7.4) that the conductivity will then be given by

$$\sigma = \frac{ne^2\Lambda}{m\,v_{\text{F}}} \qquad (8)$$

where n is the electron density, v_{F} the Fermi velocity and Λ the mean free path. For Λ we have the formula

$$\frac{1}{\Lambda} = \frac{N_i}{V}2\pi\int_0^{2\pi}(1-\cos\theta)\mathscr{T}(\theta)\sin\theta\,d\theta, \qquad (9)$$

in which $\mathscr{T}(\theta)$ is the differential scattering cross section of each of N_i impurities per unit volume.

Our assumption is, of course, that the scattering cross section is to be calculated (in Born approximation) as if it were due to a pseudo-potential $\mathscr{U}_a(\mathbf{r}-\mathbf{R})$ anchored to the ion at $\mathbf{R}$. Elementary quantum theory then gives

$$\mathscr{T}(\theta)=\frac{2\pi}{\hbar v_F}\left|\frac{1}{N}\mathscr{U}_K\right|^2\frac{3}{4}\frac{n}{\mathscr{E}_F}\frac{1}{4\pi}, \qquad (10)$$

where $\mathbf{K}$ is the scattering vector, whose magnitude, $2k_F\sin\frac{1}{2}\theta$, ranges from zero up to the diameter of the Fermi sphere. Putting (10) into (9), we may write

$$\Lambda_{\mathrm{gas}}=\frac{2\sqrt{2}}{3\pi\mathscr{E}_F^{1/2}}\left|\frac{\mathscr{E}_F}{\mathscr{U}}\right|^2, \qquad (11)$$

with

$$|\mathscr{U}|^2\equiv\frac{1}{4k_F^4}\int_0^{2k_F}|\mathscr{U}_K|^2K^3\,dK. \qquad (12)$$

If we knew the functional form of $\mathscr{U}_K$ we could evaluate these expressions. As it is, for each metal we only know a single value, $\mathscr{U}_g$, where g is a little larger than $2k_F$. But if we assume that this value is representative—that $\mathscr{U}_K$ is not rising very rapidly in the range where the main contribution is made to (12)—we can put $\mathscr{U}_g$ in place of $\mathscr{U}$ in (11), and arrive at some estimate of the mean free path. We know $\mathscr{U}_g$ independently, so here is a 'theoretical' number which is plausibly comparable with an 'observed' value of Λ, deduced, through (8), from the known electrical conductivities of the liquid alkali and noble metals. This comparison is made in Table 1.

Table 1

Metal	$\mathscr{E}_F$ (ev)	$\mathscr{U}_g$ (ev)	Λ_{gas} (Å)	Λ_{obs} (Å)	$\Lambda_{\mathrm{obs}}/\Lambda_{\mathrm{gas}}$
Li	4·76	1·53	2·2	42	19
Na	3·20	0·13	287	155	0·5
K	2·12	0·25	40·5	156	4
Rb	1·81	0·44	10·7	95	9
Cs	1·57	0·65	3·9	64	16
Cu	7·04	2·6	2·3	31	13
Ag	5·51	2·0	2·6	45	17
Au	5·51	2·0	2·6	28	11

$\mathscr{U}_g=\frac{1}{2}$ energy gap calculated by Ham and Segall.
Λ_{obs} from Cusack and Enderby (1960), and Faber (private communication).
$\mathscr{E}_F$ is the Fermi energy, in the solid at low temperatures. The values for the liquid differs from this by a few percent because of thermal expansion.

We see at once that there is a qualitative correlation between Λ_{gas} and the observed mean free path. Na and K, with small energy gaps, show long mean free paths. Li and the noble metals, with relatively large gaps, have much shorter mean free paths. The size of the pseudo-potential is obviously a relevant factor in the electrical resistance of the liquid.

§ 3. Correlation in the Liquid†

Why is there no resistance in a pure perfect crystal at zero absolute temperature, even though the electrons are certainly scattered by the ions of the crystal lattice? The reason is well understood: the ions are arranged in such regular arrays that the matrix elements are dominated by coherent interference conditions as in (2). We get modifications to the dynamical properties of the electrons, rather than incoherent random scattering with irreversible consequences.

In a liquid, also, the arrangement of the ions is by no means random. We know, for example, that two atoms may not approach closer than the sum of their radii (if they are hard spheres) and that not more than about a dozen others can be clustered about a given atom. The position vectors of the ions, $\mathbf{R}_i$, say, are not independent random variables, as we assumed in (2), but are correlated with each other in a very complicated way.

Let us put this obvious circumstance into the theory by calculating the matrix element of an electron for transitions induced by the total potential of this whole assembly of atoms. We write

$$\mathscr{U}(\mathbf{r}) = \textstyle\sum_i \mathscr{U}_a(\mathbf{r} - \mathbf{R}_i), \qquad (13)$$

and calculate as in (2):

$$\begin{aligned}\mathscr{U}_{\mathbf{k},\mathbf{k}'} &= \frac{1}{V}\textstyle\sum_i \displaystyle\int \exp[i(\mathbf{k}'-\mathbf{k})\,.\,\mathbf{r}]\mathscr{U}_a(\mathbf{r}-\mathbf{R}_i)\,d\mathbf{r}\\ &= \frac{1}{V}\textstyle\sum_i \exp(i\mathbf{K}\,.\,\mathbf{R}_i)\displaystyle\int \exp[i\mathbf{K}\,.\,(\mathbf{r}-\mathbf{R}_i)]\mathscr{U}_a(\mathbf{r}-\mathbf{R}_i)\,d\mathbf{r}\\ &= \mathscr{U}_{\mathbf{K}}\frac{1}{N}\textstyle\sum_i \exp(i\mathbf{K}\,.\,\mathbf{R}_i) \qquad (14)\end{aligned}$$

with $\mathscr{U}_{\mathbf{K}}$ exactly as in (3).

In general, this sum is a rather irregular function of $\mathbf{K}$, of more or less random phase. But we are interested mainly in the mean value of the square modulus of the matrix element, which can be expressed in terms of the radial distribution function of the atoms in the liquid. It is shown in treatises on x-ray diffraction (e.g. Fournet 1957, p. 238) that

$$\begin{aligned}a(K) &\equiv \left|\frac{1}{N}\textstyle\sum_i \exp(i\mathbf{K}\,.\,\mathbf{R}_i)\right|^2\\ &= 1 + \frac{N}{V}\int_0^\infty \{P(R)-1\}\frac{\sin(KR)}{KR}\,4\pi R^2\,dR \qquad (15)\end{aligned}$$

---

† After completing this work I was reminded by Dr. Shoenberg of the paper by Krishnan and Bhatia (1945) which anticipated the argument of this section, and which was unfortunately not known to me. Because, at that time, the idea of the pseudo-potential had not been developed, they had to guess the scattering cross section of each ion from the atomic form-factor of rare-gas atoms for slow electrons, and were unable to give a detailed account of the differences between individual metals. But they pointed out the important difference between the monovalent metals, where the Fermi level comes below the peak in the x-ray scattering function, and the polyvalent metals, where it goes beyond the peak.

where $P(R)$ is the probability of finding another atom at a distance R from a given atom in the fluid.

The next step is obviously to replace $\mathscr{U}_K$ in (10) by the function (14). This has simply the effect of introducing the function $a(K)$ into the integral, so that in place of (11) and (12) we may write

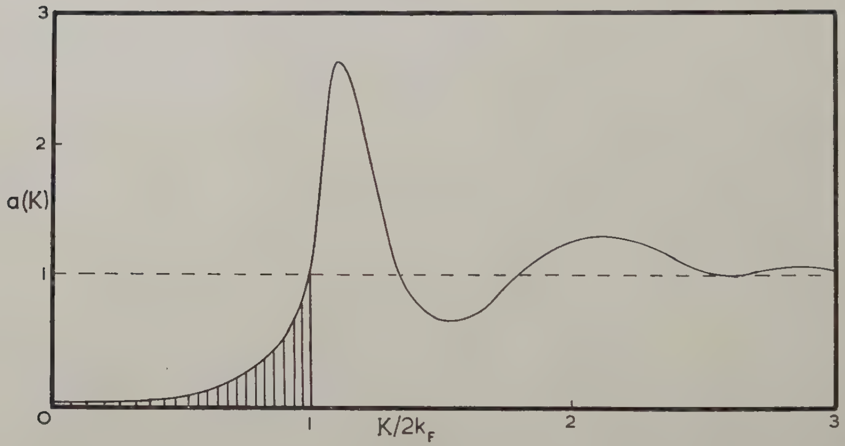
$$\Lambda_{\text{liq}} = \frac{2\sqrt{2}}{3\pi\mathscr{E}_{\mathrm{F}}^{1/2}} \frac{\mathscr{E}_{\mathrm{F}}^2}{\langle a\rangle|\mathscr{U}|^2} \quad . \quad . \quad . \quad (16)$$

where

$$\langle a\rangle|\mathscr{U}|^2 \equiv \frac{1}{4k_{\mathrm{F}}^4}\int_0^{2k_{\mathrm{F}}} |\mathscr{U}_K|^2 a(K) K^3\, dK. \quad . \quad . \quad . \quad (17)$$

That is to say, the effect of correlation is to divide Λ_{gas} by the quantity $\langle a\rangle$, defined by (15) and (17).

Fig. 1



The neutron diffraction pattern of liquid Rb, showing the typical form of $a_{\mathrm{L}}(K)$ (Gingrich and Heaton 1961).

The function $a(K)$ has been extensively studied, both experimentally and theoretically. In a gas, of course, $a(K)=1$ for all values of K. But in a true liquid it has two characteristic features, which may be seen in fig. 1.

In the first place, at a certain wavelength known to correspond approximately to the mean interatomic spacing in the fluid, there is a pronounced peak. The exact position and shape of this peak are matters for theoretical discussion, and the problem has not been resolved for fluids as dense as liquid metals, but there is no doubt that it comes from the packing of atoms around one another in the liquid in various arrangements which, if not regularly ordered over large distances, tend to be more or less uniform locally. The value of K at which the peak appears is certainly inversely proportional to the cube root of the atomic volume in the liquid ; in practice it seems to be about the same as the length of the reciprocal lattice vector for the nearest zone boundary in the solid, as one might have expected from the rules of close packing. But we know that in both alkali and noble metals the 'free-electron sphere' does not touch the zone boundary, so that $2k_F$ would not take us as far as the peak of $a(K)$ in (17). If there were more electrons per atom—say 1·3 to 1·4—we should expect important effects from the peak, but in the monovalent metals these can probably be ignored.

Looking at values of K below the peak, we observe the other significant feature; in this region $a(K)$ falls away to small values, much less than unity. This also has an obvious physical interpretation. The density of atoms in the liquid tends to a constant over large distances so that the Fourier components of long wavelength must become small.

In fact it is easy to give the limiting value of $a(K)$ at $K=0$. A simple classical analysis in which there is equipartition of energy, of amount $\boldsymbol{k}T$, in the classical degrees of freedom—the density fluctuations—gives (Ornstein and Zernike 1914)

$$a_L(0) = \boldsymbol{k}T\,N\beta_L/V, \qquad \ldots\ldots\ldots \qquad (18)$$

so that we need only know the compressibility β_L of the liquid at, say, the melting point. The data for Na are available (*Liquid Metals Handbook*, 1955); we find

$$a(0) \sim 0{\cdot}025 \qquad \ldots\ldots\ldots \qquad (19)$$

at the melting point. Thus, if this value of $a(K)$ were maintained up to $K=2k_F$, we should find $\Lambda_{\text{liq}} \sim 40\Lambda_{\text{gas}}$, which is more than enough to explain the ratios of $\Lambda_{\text{obs}}/\Lambda_{\text{gas}}$ in table 1.

But $\langle a \rangle$ is probably rather larger than 1/40. The experimental x-ray and neutron diffraction data are not available for small values of K, but in the case of Rb, for example (Gingrich and Heaton 1961), it falls rapidly from about unity at $K=2k_F$ to about 0·05 at $K=k_F$. If there were a good theoretical model we might attempt to estimate $a(K)$ from this, but such calculations as have been done (see Fournet 1957, Green 1960) are unreliable at the packing density of ordinary liquids, and we should commit the error of misplaced concreteness if we were to compute a number from any of these. There is also the unknown functional form of $\mathscr{U}_K$ to put into the integral.

All that we can say is that it may be reasonable to take

$$\langle a \rangle |\mathscr{U}|^2 \sim 0{\cdot}1 |\mathscr{U}_g|^2, \qquad \ldots\ldots\ldots \qquad (20)$$

but the number could lie between 0·05 and 0·2. In other words, the effect of correlation must be to increase the mean free path, so that Λ^{liq} might be between 5 and 20 times Λ_{gas}. (But see note on p. 1034).

§ 4. PLASMA SCATTERING

In table 1 we find the ratios of Λ_{obs} to Λ_{gas} lying between 9 and 17, except for Na and K, where they are 0·5 and 4, respectively. Thus, the correction for correlation brings the theory comfortably into the experimental range, except for these two metals where the pseudo-potential is believed to be very small. Can we account for the remaining discrepancy?

It may be, of course, that the pseudo-potential is much larger in these metals than we have assumed. But Heine (private communication) has offered a much more fruitful and plausible explanation which we shall now explore.

This argument, in its primitive form, is simply that in a liquid there are many 'holes'—as one might say, 'partial lattice vacancies' in thermal equilibrium with the surrounding atoms—which also scatter the conduction electrons. Such a 'hole' is a region deficient in positive charge, and so will, by electrostatic repulsion alone, deflect an electron. This scattering is not really included in the calculation from the pseudo-potential, and yet could be important, especially in a metal such as Na where $\mathscr{U}_{\mathbf{g}}$ is almost zero.

To make a mathematical theory of this effect, we should consider the local distribution of density in the fluid, which cannot be precisely uniform. We define the local dilatation, $\Delta(\mathbf{r})$, and ask for its mean square Fourier components, since these will occur in the matrix elements for scattering. As shown by Landau and Lifshitz (1958, p. 363), this takes us back again to the molecular distribution function:

$$\overline{|\Delta_{\mathbf{K}}|^2}=a(K). \qquad \qquad (21)$$

In particular, the mean value of $|\Delta(\mathbf{r})|^2$ is $a(0)$, as already noted in (18).

But the effect of a dilatation of amount Δ is not simply to scatter as if by a charge $n\boldsymbol{e}\Delta$. The system of ions and electrons constitutes a plasma, where fluctuations of density in the heavy ions are automatically screened by the motion of the electrons. In a degenerate Fermi plasma, however, this screening is incomplete because it would give rise to variations in the Fermi level. It can easily be shown, in a variety of ways (Ziman 1960, chap. V) that the effective potential associated with a dilatation is the 'plasma potential',

$$\mathscr{W}_0\Delta=\tfrac{2}{3}\mathscr{E}_{\mathrm{F}}\Delta. \qquad \qquad (22)$$

This is the long-wave limit of the matrix element of this potential, and, in a solid, it has been taken to be the dominant term, being at the heart of the Bardeen formula for the electron–phonon interaction. From (18), (21) and (22) we can say that the transition probability for the 'plasma scattering' is proportional to

$$|\mathscr{W}_0|^2a(0)=(\tfrac{2}{3}\mathscr{E}_{\mathrm{F}})^2\frac{\boldsymbol{k}TN\beta}{V} \qquad \qquad (23)$$

for small scattering angles—a quite precise result.

But when we try to work out $\mathscr{W}_K$ for larger values of K we run into difficulties, which have not really been resolved even in the case of a solid. Large values of K correspond to fluctuations of short wavelength, and we do not know exactly how to allow for inadequate screening. Moreover, the ions are of finite size, and cannot be treated as point charges. Abrupt variations in the apparent density of atomic centres do not give rise to such sharp variations in the scattering potential, and will be further softened by charge shifts in the electron gas.

It seems likely, however, that $\mathscr{W}_K$ is a function that must fall off quite rapidly as K increases. To fix our ideas, we might refer to the Bardeen formula, where the charge of each conduction electron is spread over a Wigner–Seitz sphere (radius r_s) containing one atomic volume. In that case, excluding a screening factor which does not vary rapidly,

$$\mathscr{W}_K \approx \tfrac{2}{3}\mathscr{E}_F \mathscr{G}(Kr_s) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (24)$$

where

$$\mathscr{G}(x) \equiv 3(x\cos x - \sin x)/x^3.$$

If we square $\mathscr{W}_K$, and average with weight $K^3 a(K)$, as in (17), we find

$$\Lambda_{\text{plasma}} = \frac{2\sqrt{2}}{3\pi \mathscr{E}_F^{1/2}} \left(\frac{\mathscr{E}_F}{\frac{2}{3}\mathscr{E}_F}\right)^2 \frac{1}{\langle a\rangle\langle \mathscr{G}^2\rangle}. \quad . \quad . \quad . \quad (25)$$

A crude estimate for $\langle \mathscr{G}^2\rangle$ is about 0·1. Thus, the plasma effect alone gives rise to a resistance in the liquid, even if there is no pseudo-potential; in the case of Na, putting numbers into (25) gives us a value between 50 and 200 Å, which is obviously of just the right order of magnitude to explain the observed mean free path.

If this argument is true, and if the other monovalent metals are similar to Na, we can use the experimental result for that metal to calibrate the plasma term in all the others. That is, if $\langle a\rangle\langle \mathscr{G}^2\rangle$ is the same for all liquid monovalent metals at their melting point—a plausible assumption—then (25) tells us that $\Lambda_{\text{plasma}}\mathscr{E}_F^{1/2}$ should be a constant†. Taking the experimental value of this constant for Na as our standard, we calculate the plasma resistance for each of the other liquid metals, and subtract it from the total, leaving a residual resistance, which we may express again as a mean free path, $\Lambda_{\text{structure}}$. Thus

$$\frac{1}{\Lambda_{\text{structure}}} = \frac{1}{\Lambda_{\text{obs}}} - \frac{1}{\Lambda_{\text{plasma}}}. \quad . \quad . \quad . \quad . \quad (26)$$

The result of this calculation is shown in table 2.

Table 2

	Li	Na	K	Rb	Cs	Cu	Ag	Au
$\Lambda_{\text{structure}}$ (Å)	62		805	175	89	44	72	37
Λ_{gas} (Å)	2·2	—	40·5	10·7	3·9	2·3	2·6	2·6
$\dfrac{\Lambda_{\text{structure}}}{\Lambda_{\text{gas}}}$	28	—	20	16	23	19	28	14

---

† This is equivalent to the rough theoretical rule that the mean free path at the melting point in a free-electron solid metal should be a fixed multiple of the lattice spacing—see Ziman 1960, § 9.7.

It is evident that the value of $\Lambda_{\text{structure}}$ now fits excellently into our theory; in each case it is about 20 times the value of Λ_{gas} deduced from the pseudo-potential, as if the effect of correlation in the liquid were to make $1/\langle a \rangle \sim 20$. There is no systematic trend left in this table; the different numbers in the final row could easily reflect errors in our estimates of $\mathscr{U}_{\mathbf{g}}$ and of Λ_{obs}. We know, for example, that the 'necks' of the Fermi surface of Au are rather larger than they are for Ag, suggesting a rather larger value of $\mathscr{U}_{\mathbf{g}}$ in Au (see Ziman 1961), whereas the data of tables 1 and 2 come from Segall's direct calculations, which show '$\sim 0{\cdot}3$ rydbergs' for both metals.

§ 5. The Displacement Potential

Formally, eqns. (16), (25) and (26) may be combined. We may write for the estimated mean free path

$$\frac{1}{\Lambda} = \frac{3\pi \mathscr{E}_{\mathrm{F}}^{1/2}}{2\sqrt{2}} \frac{\langle C^2 \rangle \langle a \rangle}{\mathscr{E}_{\mathrm{F}}^{2}}, \qquad (27)$$

where we have defined a new function

$$|C(K)|^2 \equiv |\mathscr{U}_K|^2 + |\mathscr{W}_K|^2 \qquad (28)$$

whose average is expressed by

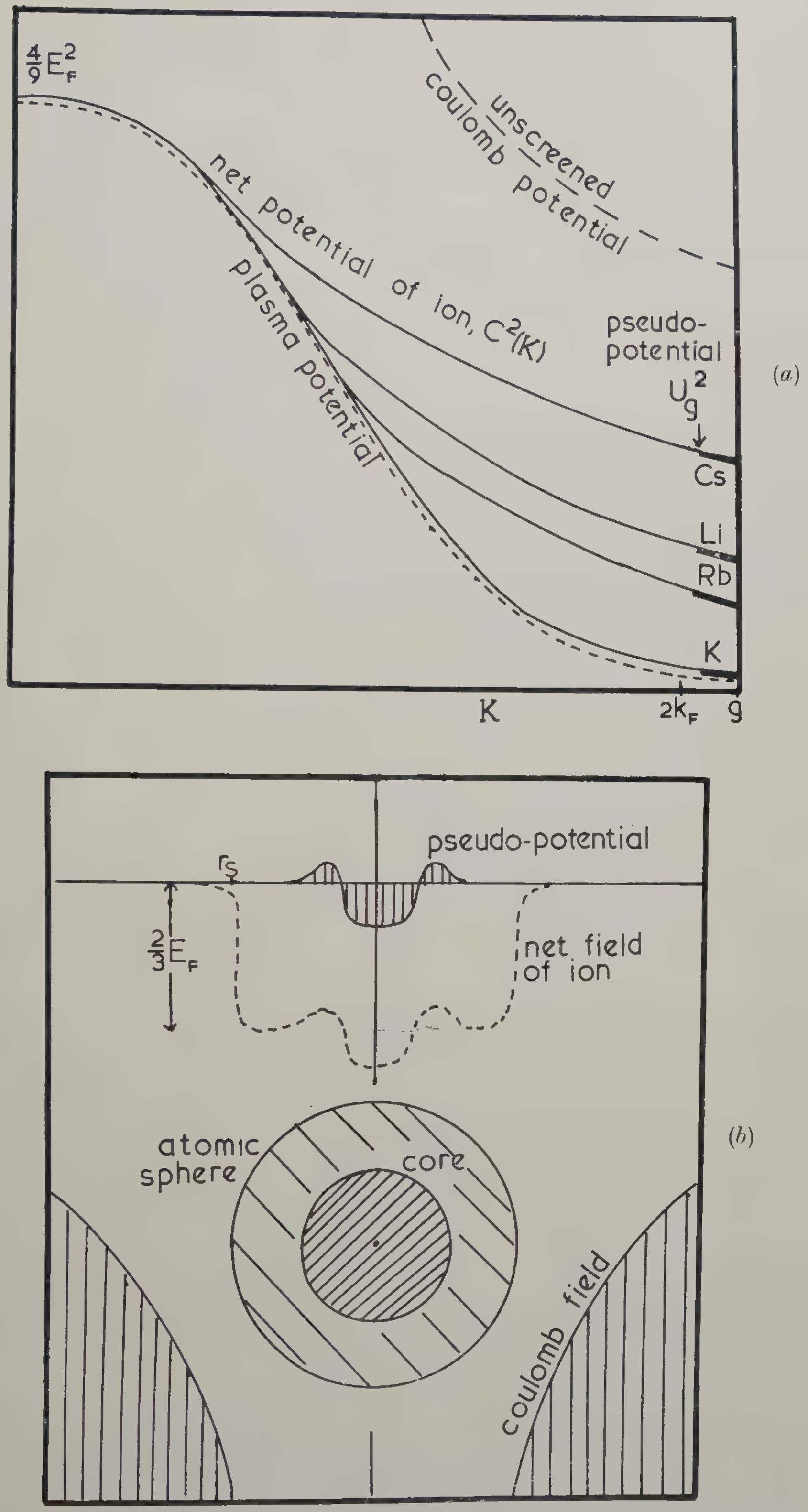
$$\langle C^2 \rangle \langle a \rangle \equiv \frac{1}{4k_{\mathrm{F}}^4} \int_0^{2k_{\mathrm{F}}} |C(K)|^2 a(K) K^3 \, dK. \qquad (29)$$

As it stands, we have gained nothing. But by talking about $C(K)$, instead of its two separate parts, we avoid the suggestion that pseudo-potential and plasma scattering are essentially distinct mechanisms. They both arise from the effective field which the displacement of an ion exerts upon a conduction electron, and are different only in the range of values of K for which they are important. For large values of K we have what we have called the pseudo-potential, which arises as the net difference between Coulomb attraction and Pauli repulsion in the ion core. Cohen and Heine (1961) have shown that it is possible to achieve some exact cancellations between these effects, particularly for the Fourier components of very large wave number. Thus one can eliminate the deep potential inside the ion core. Nevertheless, the pseudo-potential is still probably relatively localized, since we know that it has large Fourier components for waves like (111) or (200), and there is no reason why $\mathscr{U}_K$ should decrease abruptly beyond this point.

On the other hand, going to smaller values of K corresponds to studying the potential of the ion at larger distances from its nucleus, and this, we know, must be just a simple Coulomb potential, $\mathbf{e}^2/r$. There is no orthogonalization in this region, so that this part of the potential is not cancelled. In k-space this means a singularity like $1/K^2$ as $K \to 0$. But this singularity (which would give an infinite scattering cross section, etc.) is screened by the plasma of conduction electrons.

In principle we should solve the problem of the collective motion of the electrons around the ion, but in practice the result is the same as if the ionic potential had the Fourier components of $\mathscr{W}_K$, as in (24), say. In fact

Fig. 2



(*a*) The square of the matrix element of the scattering potential, showing the Bardeen form of the plasma potential, the values of the pseudo-potentials at the zone boundary, and the general form of the function $C^2(K)$ that is assumed to connect them.

(*b*) The field of an ion, indicating the general form of the net field combining the pseudo-potential in the core with a screened coulomb field in the Wigner–Seitz cell.

we should have obtained exactly this result if we had supposed (quite plausibly) the electrostatic field of the ion to be a square-well potential, of depth $\frac{2}{3}\mathscr{E}_{\mathrm{F}}$, extending out to the radius, r_{S}, of the Wigner–Seitz cell which it dominates, but not to be observed in the next cell. Instead of the infinity at $K=0$, we get exactly the Fermi–Thomas screening potential, $\frac{2}{3}\mathscr{E}_{\mathrm{F}}$. One can even show (see Ziman 1960, § 5.7) that any part of the 'ordinary' pseudo-potential at small K will be screened by the plasma electrons, and will not need to be added separately to the plasma potential.

The picture now is extremely simple. For all the monovalent metals $C(K)$ for small K is equal to $\frac{2}{3}\mathscr{E}_{\mathrm{F}}$. From this point it falls fairly steeply; for Na, it is effectively zero by the time we reach $K=2k_{\mathrm{F}}$. But in the other metals the profile is more complex, with a shoulder and a tail reaching out to larger wave numbers. The height of this shoulder is measured by $\mathscr{U}_{\mathbf{g}}$—our estimate of the pseudo-potential at the zone-boundary. Because $a(K)$ rises fairly fast as we approach $2k_{\mathrm{F}}$, the resistance of the liquid tends to be dominated by the pseudo-potential, even though elsewhere the plasma part has larger Fourier components. The shapes of $\mathscr{U}_K$ and of $\mathscr{W}_K$ are such that they are important over different ranges of K, so that there is no essential difference between the sum of their squares (as in (28)) and the square of their sum. That is why we can treat them approximately as if they contributed independently to the total resistance.

§ 6. CHANGE OF RESISTANCE ON MELTING

A quantity which can be measured fairly easily, and which being dimensionless, invites theoretical speculation, is the ratio $\rho_{\mathrm{L}}/\rho_{\mathrm{S}}$, of the electrical resistivities of the liquid and solid phases in the neighbourhood of the melting point. This ratio is set out in table 3, along with a parameter expressing directly the strength of the pseudo-potential in units of the Fermi energy.

Table 3. Resistance change on melting

	Li	Na	K	Rb	Cs	Cu	Ag	Au
$\rho_{\mathrm{L}}/\rho_{\mathrm{S}}$	1·68	1·45	1·56	1·60	1·66	2·04	2·09	2·28
$\mathscr{U}_{\mathbf{g}}/\mathscr{E}_{\mathrm{F}}$	0·32	0·04	0·12	0·24	0·41	0·37	0·36	0·36

It is obvious that these two parameters are strongly correlated. In the noble metals there is a significant volume expansion on melting, which would upset an exact correspondence, but amongst the alkali metals the correlation is excellent†. The increase of resistance on melting is bigger in metals where the pseudo-potential is greater.

---

† We can even guess that the estimate of $\mathscr{U}_{\mathbf{g}}$ for Li is a bit low; it should be as high in proportion, as in Cs—which would bring the ratio for Li in table 2 almost exactly up to the average. Between Ag and Au, also, the difference in $\rho_{\mathrm{L}}/\rho_{\mathrm{S}}$ supports our interpretation of the differences in table 2, where probably the theoretical estimate of $\mathscr{U}_{\mathbf{g}}$ for Ag is too large.

There exists a detailed theory of the resistivity of the solid metals (Ziman 1960, Collins and Ziman 1961) which fits the facts fairly well. But this theory is expressed in different terms from that for the liquid, and a direct quotient of the two formula is not easy to analyse. Let us, instead, try to calculate the resistivity of a solid along the same lines as for liquid. That is, in both phases, we shall assume that the same function, $C(K)$, represents the effective potential of an ion for scattering electrons.

Now this means that in (27) and (29), say, all that is altered is the correlation function $a(K)$. For a solid this can be calculated from first principles through the definition (15);

$$a_{\mathrm{S}}(\mathbf{K}) = \left| \frac{1}{N} \sum_l \exp(i\mathbf{K} \,.\, \mathbf{R}_l) \right|^2 . \qquad \ldots\ldots \quad (30)$$

The position of the ion, which would in a perfect lattice have been at $\boldsymbol{l}$, is given by

$$\mathbf{R}_l = \boldsymbol{l} + \sum_{\mathbf{q}}{}' \mathbf{y}_{\mathbf{q}} \exp(-i\mathbf{q} \,.\, \boldsymbol{l}); \qquad \ldots\ldots \quad (31)$$

it is displaced by the lattice waves, whose wave-vectors are $\mathbf{q}$ and amplitudes $\mathbf{y}_{\mathbf{q}}$. If the $\mathbf{y}_{\mathbf{q}}$ are small, we can expand the potential in (30) about $\boldsymbol{l}$, and have

$$\begin{aligned} a_{\mathrm{S}}(\mathbf{K}) &= \left| \frac{1}{N} \sum_l \{\exp(i\mathbf{K} \,.\, \boldsymbol{l})[1 + i\mathbf{K} \,.\, \textstyle\sum_{\mathbf{q}} \exp(-i\mathbf{q} \,.\, \boldsymbol{l})]\} \right|^2 \\ &= |\delta_{\mathbf{K},\mathbf{g}} + i\textstyle\sum_{\mathbf{q}}(\mathbf{y}_{\mathbf{q}} \,.\, \mathbf{K})\delta_{\mathbf{K}-\mathbf{q},\mathbf{g}}|^2 \\ &= \delta_{\mathbf{K},\mathbf{g}} + |\mathbf{K} \,.\, \mathbf{y}_{\mathbf{K}-\mathbf{g}}|^2 . \qquad \ldots\ldots \quad (32) \end{aligned}$$

The initial delta-function need not concern us; it just takes us back to (2) and the band structure of the metal. But the other term shows us that $a_{\mathrm{S}}(K)$ is proportional to the square of the amplitude of the lattice vibrations of wave-vector $\mathbf{K}-\mathbf{g}$. If $\mathbf{K}$ lies within the first Brillouin zone, this is the wave with $\mathbf{q}=\mathbf{K}$; otherwise, it is the phonon whose wave-vector is just the value of $\mathbf{K}$ 'reduced' into that zone. In other words, we are dealing with electron–phonon N-processes and U-processes, as in the conventional theory.

At the melting point we can treat the thermal distribution of lattice modes as if it were classical; if the frequency of the mode is $\nu_{\mathbf{K}-\mathbf{g}}$, then

$$\overline{|\mathbf{y}_{\mathbf{K}-\mathbf{g}}|^2} = \boldsymbol{k}T/M\nu_{\mathbf{K}-\mathbf{g}}{}^2 \qquad \ldots\ldots \quad (34)$$

where M is the mass of an ion. When $\mathbf{K}$ lies in the first zone, we need only consider longitudinal modes, whose velocity, in an isotropic (polycrystalline) solid, is given directly from the compressibility, β_{S}. Thus, we find

$$a_{\mathrm{S}}(0) = \boldsymbol{k}TN\beta_{\mathrm{S}}/V, \qquad \ldots\ldots \quad (35)$$

as one might expect by analogy with (18).

This formula should hold for small values of K, up to the point where the dispersion of the longitudinal waves becomes apparent—i.e. as K

approaches the zone boundary (or, in the Debye approximation, the boundary of the Debye sphere). Thereafter one finds considerable complications, all of which tend to increase $a_S(K)$. Thus:

(i) K goes on increasing, whereas $|\mathbf{K}-\mathbf{g}|$ begins to decrease, and ν with it. Indeed, there are singular points ($\mathbf{K}=\mathbf{g}$) where $\nu \to 0$ again, whilst K is still finite. In other words, the poles of the delta-function in (32) are echoed, and broadened, by the thermal term.

(ii) $\mathbf{K}$ need not be parallel to $\mathbf{K}-\mathbf{g}$, so that there will be contributions from transverse modes. These have a lower velocity than the longitudinal modes, so they will, again, increase $a_S(K)$.

The complete calculation of $a_S(K)$ is thus possible in principle. According to Gerstenkorn (1952), and to Bhatia and Krishnan (1948) who have made serious attempts to give numerical values to this function, the effect of transverse waves is very important; they find $a_S(K)$ rising by a factor of 50 as K goes from 0 to $2k_F$. But if one then includes in the calculation a Debye–Waller factor (which allows for the error in the approximation (32) when the terms have large amplitudes, see Bailyn 1961), one finds that $a_S(K)$ lies somewhat below the observed values of $a_L(K)$ for all values of K up to $2k_F$. This is, surely, inevitable, since the atoms in the solid, even near the melting point, are more strongly correlated in position than they are in the liquid (fig. 3).

The change on melting thus reflects a general increase of $a(K)$. This is the argument of Gerstenkorn, who showed that the theoretical change was of the right order of magnitude and calculated that ρ_L/ρ_S should lie between 1·6 and 2. But he was unable to explain the variations in the observed ratio from one metal to another, except in terms of variations in the 'number of free electrons per atom', and he probably over-estimated the effect of the transverse modes (Collins and Ziman 1961) (see note on p. 1034).

With our more realistic formulae for the scattering potential the situation becomes quite clear. Consider Na, where plasma effects are dominant, and $C(K)$ is concentrated at small values of K. The change of resistance comes from the change of $a(K)$ near $K=0$. In other words, in first approximation

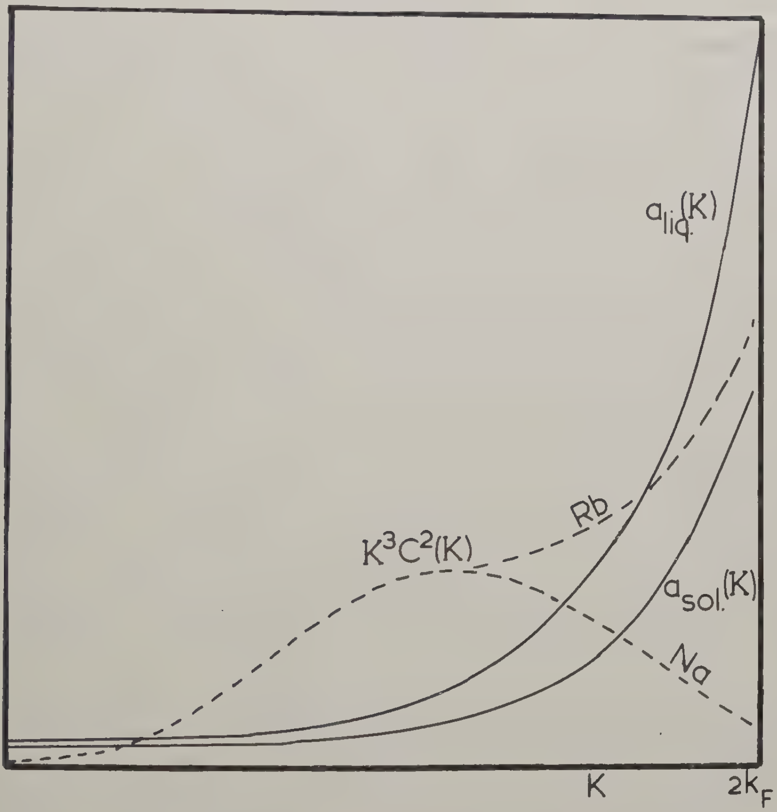
$$\frac{\rho_L}{\rho_S} \sim \frac{a_L(0)}{a_S(0)} = \frac{\beta_L}{\beta_S}. \qquad \qquad (36)$$

The ratio of the compressibilities is about 1·3; whilst the resistance ratio, in table 3, is 1·45. If, as one expects, and as Gerstenkorn calculates, $a_L(K)$ rises a little more rapidly than $a_S(K)$, this would be enough to explain the discrepancy.

In the other metals the resistance is dominated by the pseudo-potential, which becomes effective at larger values of K. But there $a_L(K)/a_S(K)$ has increased somewhat above the limit (36), to as much as 2 at some points in Gerstenkorn's table. The resistance ratio should thus increase as these larger ratios are weighted in the integral by large values of $\mathscr{U}_{\mathbf{K}}$. This is just what is observed.

We can even interpret the larger values of the resistance ratio in the noble metals, which might be expected to lie nearer to Li or Cs. In the noble metals the transverse waves do not have so low a velocity, compared with the longitudinal waves, and the elastic anisotropy is less pronounced, so that $a_S(K)$ is probably somewhat smaller, at large values of K, than it is in the alkali metals.

Fig. 3



A guess at the change of $a(K)$ on melting. The resistance is the integral of the product of $a(K)$ with $C^2(K)$, here indicated for Na (plasma potential only) and for a metal such as Rb with a large pseudo-potential.

The argument of this section cannot, however, be the whole story. In the conventional theory of the electron–phonon interaction in the solid we only have the plasma term in $C(K)$. It is true that the full Bardeen theory contains a term $\{\mathscr{U}(r_s) - \mathscr{E}_0\}$, which depends on the detailed solution of the Schrödinger equation inside a Wigner–Seitz cell, and which takes over at large values of K. This term is obviously an approximation to the pseudo-potential, but has usually been taken to be rather small. On the other hand, as shown by Collins (1961), there can be large contributions to what is, effectively, the function $C(K)$ from mixing of plane waves in the matrix elements for transitions between states near the zone boundary.

The amount of such mixing will depend on the pseudo-potential, so that, for reasons that are not the same in detail as we have assumed in the liquid, the resistance of the solid is fairly sensitive to the size of the energy gap, etc. For the moment it is a moot point whether this is just another way of calculating the same effect, or whether the mixing of states is an additional effect in the solid, so that, for example

$$\rho_{\mathrm{L}}/\rho_{\mathrm{S}} < \langle a_{\mathrm{L}} \rangle / \langle a_{\mathrm{S}} \rangle, \qquad \qquad (37)$$

or whether there is, in fact, a similar mechanism at work in the liquid—an incipient band structure which also tends to mix together different plane waves (as in (4)) when we approach the equivalent of the zone boundary.

§ 7. DEPENDENCE OF RESISTANCE ON TEMPERATURE AND PRESSURE

It is interesting to consider, also, the temperature dependence of ρ_{L}. According to (27), this must come from $a(K)$, since $C(K)$ is a function only of the ionic potential, electron density, etc. Now we know from (18) that $a_{\mathrm{L}}(K)$ is proportional to the absolute temperature at small wave numbers. But the resistance of Na, coming entirely from the plasma scattering, is dominated by this part of $a_{\mathrm{L}}(K)$. Thus, the resistance of liquid Na should be nearly proportional to the absolute temperature, exactly as observed.

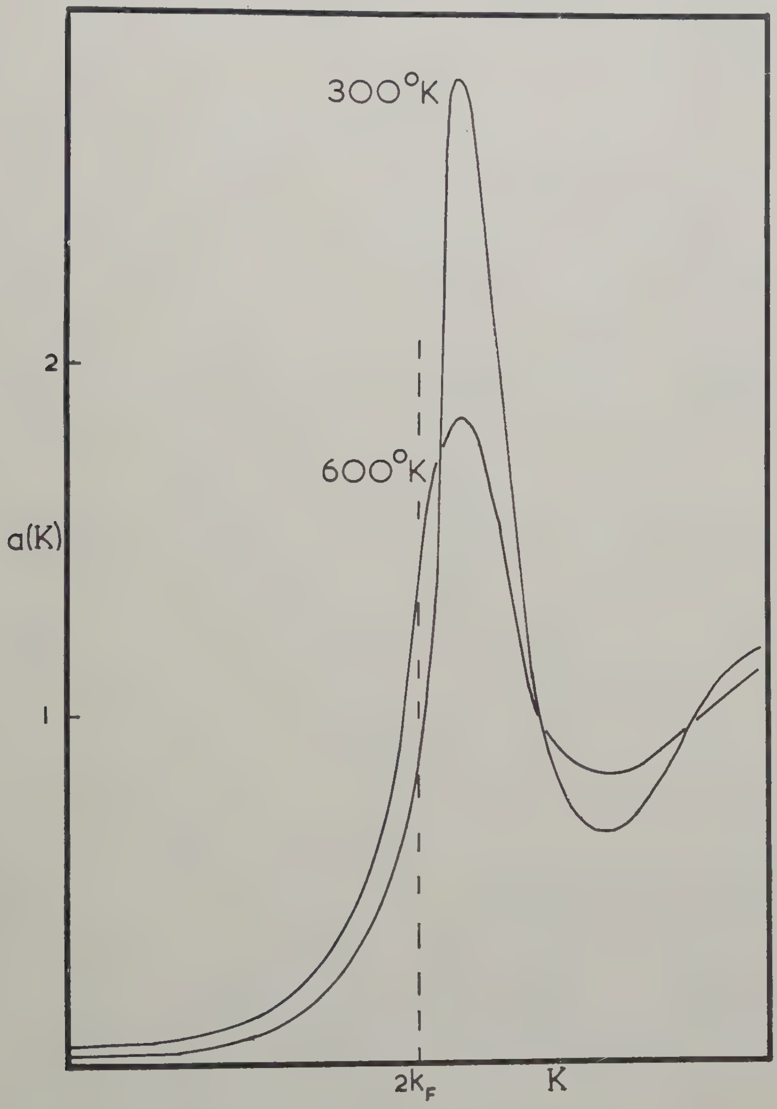
But where the pseudo-potential is important we must look at the temperature dependence of $a_{\mathrm{L}}(K)$ near $K=2k_{\mathrm{F}}$. This is not subject to any simple theory, but the observational data show (see, e.g. Gingrich 1943, Gingrich and Heaton 1961) that here a_{L} increases only slowly with temperature (fig. 4). After all, if the density of the fluid is kept constant (we assume that we are studying $\rho_{\mathrm{L}}(T)$ at constant volume), there is little room for large alteration in the local radial distribution function, $P(R)$,

Table 4. Temperature and volume coefficients of resistivity at the melting point, T_{m}. $\rho_0=\rho_{\mathrm{L}}-T_{\mathrm{m}}(\partial\rho_{\mathrm{L}}/\partial T)_V$; the so-called 'residual' resistance. The data in this table are subject to considerable error in some cases

	Li	Na	K	Rb	Cs	Cu	Ag	Au
ρ_{L} ($\mu\Omega$ cm)	24	9·6	13	23	37	21	17	31
$T_{\mathrm{m}}\left(\dfrac{\partial\rho_{\mathrm{L}}}{\partial T}\right)_P$ ($\mu\Omega$ cm)	14	12	20	25	33	12	11	19
$\dfrac{V}{\rho_{\mathrm{L}}}\left(\dfrac{\partial\rho_{\mathrm{L}}}{\partial V}\right)_T$	<0	4·35	4·75	3·95	1·8	—	—	—
$T_{\mathrm{m}}\left(\dfrac{\partial\rho_{\mathrm{L}}}{\partial T}\right)_V$ ($\mu\Omega$ cm)	16	8·0	14	15	27	—	—	—
ρ_0 ($\mu\Omega$ cm)	8	1·6	-1	8	10	—	—	—

which occurs in (15). Thus, the 'structural' resistance need not be proportional to T, although it should probably increase slowly with temperature, at least in the monovalent metals.

Fig. 4



The change of $a_{\mathrm{L}}(K)$ on doubling the absolute temperature. Curves again sketched from neutron-diffraction pattern for Rb (Gingrich and Heaton 1961).

Such data as are available in the literature have been collected and analysed by Faber (private communication) and are set out in table 4. These show that, at constant volume, ρ_{L} is proportional to T only for

Na and K. In the other alkali metals there seems to be a 'residual resistance', ρ_0, which makes up about one-third of ρ_L at the melting point and which stays constant whilst the remainder increases proportionately with T. This term has been attributed to scattering by the disordered structure of the liquid (Cusack and Enderby 1960) but, if we compare its magnitude with the 'structural resistance' in our analysis —the part of ρ_L associated with the pseudo-potential—we find that ρ_0 is rather too small. In other words, the 'structural resistance' is not a constant but increases with T, although at a much slower rate than the plasma term. This is exactly what our theory predicts.

The variation of ρ_L with volume is also qualitatively in agreement with our model. The data for the alkali metals in table 4 show that Na and K have $(V/\rho_L)(\partial\rho_L/\partial V)_T$ between $+4$ and $+5$ at the melting point, whereas the coefficient falls to $+1{\cdot}8$ in Cs and is between 0 and -1 in Li. It seems as if the volume coefficient for 'plasma resistance' is about 4, whilst the 'structural resistance' is rather less sensitive to changes of volume.

In theory the plasma scattering ought to depend mainly on $\partial a(0)/\partial V$ which, in turn, depends mainly on the variation with volume of the compressibility. Both in liquid and solid alkali metals this is probably determined by the properties of the gas of conduction electrons, and it is not surprising that the compressibility decreases when the volume is decreased. Standard theory for the solid (e.g. Ziman 1960, § 9.15) calculates the volume coefficient of resistance as twice the Grüneisen constant, so that a value around $+4$ is reasonable. It is interesting that, in Na, $(V/\rho)(\partial\rho/\partial V)_T$ does not change very much on melting.

On the other hand, the structure scattering is a much more complicated effect. The pseudo-potential may be affected by compression, and there are compensating scaling corrections to be made on $2k_F$ and on $a(K)$. But the general effect can be judged from the fact that compression of the liquid tends to 'sharpen' the peak of the x-ray scattering function, having much the same effect as a reduction of temperature (fig. 4). In the neighbourhood of $2k_F$, $a(K)$ may well be raised, rather than lowered, by compression, and $\partial a(K)/\partial V$ would then be negative. The volume coefficient for structure scattering is thus expected to be small, if positive, and could easily be negative.

We might here consider Mott's theory (1934) of the resistance of liquid metals. This was worked out for an 'Einstein' model, with the ion oscillating about its 'lattice site' with frequency ν_E. In other words, it is a 'gas' approximation, in which the contributions of the ions are summed independently. In the solid (see Jones 1956) this approximation is plausible, up to the point of giving a linear temperature dependence at high temperatures. But even then, it is somewhat spurious. One can easily show that it is equivalent to writing (27) with (29), with

$$a(K) = \frac{\mathbf{k}T}{M\nu_E^2} K^2, \qquad \ldots\ldots\ldots \quad (38)$$

which does not give proper weight to those processes where a long-wave phonon is involved. This is particularly unfortunate for the U-processes, which contribute, we know, a great part of the resistance. We can, of course, re-adjust the theory by a convenient choice of $C(K)$ but this is only a fudge.

But the essence of Mott's argument is that the increased resistance of a liquid may be ascribed to the increased amplitude of the vibrations of the ion about their mean positions, and that the entropy of melting may be traced to the same origin. On this model, or any simple extension of it which ignores the contribution of disorder to the resistance, ρ_L would be proportional to T. Our work shows that for metals with a large pseudo-potential, this cannot be true.

§ 8. THERMOELECTRIC POWER

Finally, let us look at the thermoelectric power Q. According to Cusack and Enderby (1960), this behaves as shown in table 5. For comparison with theory we give also the dimensionless parameter†

$$\xi = Q/Q_F = Q \bigg/ \left(\frac{\pi^2}{3}\frac{\boldsymbol{k}^2 T}{\boldsymbol{e}\mathscr{E}_F}\right) \quad . \quad . \quad . \quad . \quad (39)$$

which in theory (Ziman 1960, § 9.11) is deducible from the formula

$$\xi = \mathscr{E}_F \left[\frac{\partial \ln \sigma(\mathscr{E})}{\partial \mathscr{E}}\right]_{\mathscr{E}-\mathscr{E}_F} . \quad . \quad . \quad . \quad . \quad (40)$$

Table 5. Thermoelectric power of solid and liquid metals

	Li	Na	K	Rb	Cs	Cu	Ag	Au
Q_S (μv/deg)	15·5	−9·9	−17·8	−10·9	0·6	7·2	10·4	3·7
Q_L (μv/deg)	21·5	−9·9	−15·5	−7·1	6·4	16·7	10·4	3·7
ξ_S	−6·7	3·5	4·6	2·6	−0·12	−1·6	−1·9	−0·63
ξ_L	−9·3	3·5	4	1·7	−1·36	−3·6	−1·9	−0·63
$\xi_L - \xi_S$	−2·6		−0·6	−0·9	−1·2	−2·0	0	0

Here again we find three significant correlations with the magnitude of the pseudo-potential. (i) ξ_S is positive for Na, and becomes more negative as we go towards Li, or Cs, or the noble metals. (ii) ξ_L is positive for Na, and also becomes more negative in metals with larger energy gaps. (iii) The change of ξ on melting is zero for Na, but is negative, and of increasing magnitude, as we go out to the other metals, in the same order (except for Ag and Au, whose thermoelectric power at high temperatures in the solid state is distinctly anomalous—see Ziman 1961, § 11).

These effects can all be explained.

(i) As we see from (40), what matters is the way in which the resistance of a hypothetical metal would vary as we changed the Fermi level. This depends, in turn, on the behaviour of the scattering as k_F increases. A value around 3 has been calculated for a free-electron solid assuming the Bardeen form of matrix element (see Ziman 1960, § 9.11)—i.e. where the plasma potential is dominant—so that most weight comes from values of $a_S(K)$ where K is not too near $2k_F$. This gives us Na and, probably, potassium.

---

† The sign of ξ is opposite to that of Q, because $\boldsymbol{e}$ is taken to be negative—the charge of an electron. Thus, if ξ is positive we have ' normal ', ' electron-like ', thermoelectric power.

But when the pseudo-potential is large the resistance is determined by scattering vectors near $2k_F$—and in this region $a_S(K)$ is rising rapidly (see fig. 3). Negative values of ξ can then easily appear, simply by the great increase of scattering that can occur when K is increased slightly. In conventional terms, the U-processes are very sensitive to the size of the Fermi surface, and increase rapidly as $\mathscr{E}_F$ grows (Klemens 1960, Ziman 1961).

(ii) The same argument holds in the liquid, except that we consider $a_L(K)$ instead of $a_S(K)$.

(iii) As discussed in § 6, it seems that, near $2k_F$, $a(K)$ rises more steeply in the liquid than in the solid. This would make no difference in Na, where only small values of K are effective, but obviously makes ξ_L more negative than ξ_S when the effect of the ion core is large.

There are some interesting corollaries to this theory—for example, that the thermoelectric power in a liquid metal such as Li or Cs need not be proportional to the absolute temperature. At higher temperatures we could get a shift of relative weight from 'structure' scattering to 'plasma' scattering, which should make ξ tend back towards $+3$. It also looks as if ξ_L should be very sensitive to the addition of extra electrons, raising the Fermi level, by alloying, since this would take us even higher towards the peak of $a_L(K)$. But we shall reserve discussion of liquid alloys and polyvalent metals for another occasion.

§ 9. CONCLUSION

The theory of this paper can be summarized in words as follows: The electrical resistivity of a liquid metal is the sum of two terms, a 'thermal scattering' term and a 'structure scattering' term. The 'thermal' term is proportional to the absolute temperature. Its contribution to the thermoelectric power is 'normal' (i.e. 'electron-like'), and it decreases by about 30% on freezing. This term has nearly the same magnitude in all the monovalent metals. The 'structure scattering' term depends very much on the metal, being large in metals with strongly distorted Fermi surfaces. It is nearly independent of temperature, contributes 'anomalously' to the thermoelectric power, and ought not to be present in the solid.

A description of this sort is implicit in many discussions of the problem (e.g. Cusack and Enderby 1960). As we have seen, Mott (1934) long ago tried to relate the 'thermal' term, assumed to be the only one, to the entropy of melting, and Gubanov (1956, 1957) and Knight *et al.* (1959) have recently attempted an analysis of the 'structure scattering' as if there were, in the liquid, a local atomic arrangement similar to the solid lattice, but always changing its axes. A model like this could certainly produce a structural resistance proportional to the square of the pseudo-potential (through distortion of the Fermi surface, and mismatch of waves travelling from one locality to another). But it is difficult to make precise and is probably wrong in principle. If one accepts the work of Bernal (1959, 1960), the local 'order' in a liquid is not the same as in the solid; it is never so geometrically precise, even on the scale of a few dozen atoms, as to give such well-defined objects as polyhedral Brillouin zones.

It is gratifying, therefore, to see how by a simple perturbation treatment, this description can be made almost quantitative. The key as in Krishnan and Bhatia's treatment, is the constructive interference of waves diffracted by adjacent ions, and the mathematical link with the radial distribution function. The whole problem is characterized by two functions: $C(K)$, which is a Fourier transform of the potential of each ion after allowing for the 'other' electrons, inside and outside the core; and $a(K)$ which is just what is measured in X-ray or neutron diffraction, and which is a Fourier transform of the correlation of atomic positions.

The properties of these functions, and the behaviour of the integral of their product, allow us to discuss the following phenomena, with semi-quantitative estimates for some of the parameters:

(*a*) the electrical resistivity of liquid Na;
(*b*) the electrical resistivity of all the other monovalent metals, including the noble metals;
(*c*) the change of resistance on melting;
(*d*) the temperature dependence of the resistivity;
(*e*) the thermoelectric power, and its change on melting.

All that we put into the theory are atomic volumes (i.e. values of $\mathscr{E}_F$) independently calculated values of energy gaps and the experimental distribution functions. In every case, theory and fact fit easily without need for additional special hypotheses. We can scarcely ask for more.

Yet there remain some queries. To what extent can the pseudo-potential be written as a simple function of $\mathbf{r}$? Can we prove in detail that $C(K)$ must fall rapidly from $\frac{2}{3}\mathscr{E}_F$? When the pseudo-potential is large, are there any incipient 'band gap' effects, such as occur in the solid phase of the noble metals? How does $a_L(K)$ behave, exactly, and what is its temperature-dependence at large K? Is the conventional calculation for the resistivity of the solid consistent with our present theory? Can we carry over the argument to polyvalent metals? What happens in alloys? Can we do better than simple perturbation theory when the electrons are more tightly bound to their atoms? This paper must be regarded as only a first step towards a complete theory.

ACKNOWLEDGMENTS

It is obvious that Dr. Heine's suggestion (see § 4) was of great value in clarifying the position. I am very grateful to him for this, and many other stimulating discussions. Professor Mott, Dr. T. E. Faber, Mr. E. G. Wilson and Mr. C. C. Bradley have been most helpful with comments and suggestions.

I also wish to thank Dr. B. Segall, of the General Electric Research Laboratories, Schenectady, N.Y., U.S.A., for telling me the results of his calculations and Dr. N. Cusack, of Birkbeck College, London, for some experimental results.

Notes added in proof.—§ 3. The estimate, 0·1, for $\langle a \rangle$ now seems too small. Something more like 0·4 is indicated by the best x-ray and neutron data. This spoils the semi-quantitative agreement between theory and experiment for the magnitude of the resistivity, but the proportionality to the square of the calculated energy gaps remains. This point will be discussed in detail in a second paper dealing with polyvalent liquid metals.

§ 5. The 'displacement potential' introduced in this section turns out to be equivalent to the 'effective potential' of each ion, as calculated by Cohen and Phillips (to be published) after allowing for screening by the electron gas.

§ 6. Dr. Walter Marshall has pointed out (private communication) that Gerstenkorn is not correct in including a Debye–Waller factor in the calculation of $a_S(\mathbf{K})$. On the other hand, he probably overestimates the effect of transverse waves by ignoring the polarization factor in the electron–phonon interaction. The argument of this section needs further study to be made quantitative.

References

Bailyn, M., 1961, *Phys. Rev.*, **121,** 1336.

Bernal, J. D., 1959, *Nature, Lond.*, **183,** 141; 1960, *Ibid.*, **185,** 68; **188,** 908.

Bhatia, A. B., and Krishnan, K. S., 1948, *Proc. roy. Soc.* A, **194,** 185.

Cohen, M. H., and Heine, V., 1958, *Advanc. Phys.*, **7,** 395; 1961, *Phys. Rev.*, **122,** 1821.

Collins, J. G., 1961, *Proc. roy. Soc.* A (to appear).

Collins, J. G., and Ziman, J. M., 1961, *Proc. roy. Soc.* A (to appear).

Cusack, N., and Enderby, J. E., 1960, *Proc. phys. Soc., Lond.*, **75,** 395.

Fournet, G., 1957, *Handb. Phys.*, **32,** 238.

Gerstenkorn, H., 1952, *Ann. Phys.*, **10,** 49.

Gingrich, N. S., 1943, *Rev. mod. Phys.*, **15,** 90.

Gingrich, N. S., and Heaton, L., 1961, *J. chem. Phys.*, **34,** 873.

Green, H. S., 1960, *Handb. Phys.*, **10,** 1.

Gubanov, A. I., 1956, *Soviet Physics* (*Tech. Phys.*), **1,** 1605; 1957, *Ibid.*, **2,** 1.

Ham, F. S., 1960, *The Fermi Surface*, Ed. Harrison and Webb (New York: Wiley), p. 9.

Jones, H., 1956, *Handb. Phys.*, **19,** 227.

Klemens, P. G., 1960, *Aust. J. Phys.*, **13,** 230.

Knight, W. D., Berger, A. G., and Heine, V., 1959, *Ann. Phys.*, **8,** 173.

Krishnan, K. S., and Bhatia, A. B, 1945, *Nature, Lond.*, **156,** 503.

Landau, L. D., and Lifshitz, E..M., 1958, *Statistical Physics* (London: Pergamon).

Liquid Metals Handbook, 1955, Ed. Jackson (Washington: Atomic Energy Commission and U.S. Navy).

Mott, N. F., 1934, *Proc. roy. Soc.* A, **146,** 465.

Ornstein, L. S., and Zernike, F., 1914, *Amst. Proc.*, **17,** 793.

Phillips, J. C., and Kleinman, L., 1959, *Phys. Rev.*, **116,** 287.

Ziman, J. M., 1960, *Electrons and Phonons* (Oxford: Clarendon Press); 1961, *Advanc. Phys.*, **10,** 1.

The Direct Observation of Lattice Defects by Field Ion Microscopy

By D. G. Brandon and M. Wald†
Department of Metallurgy, University of Cambridge

[Received May 11, 1961]

Abstract

Observations of Mo and W in the field ion microscope have revealed the structure of edge dislocations and a grain boundary in molybdenum. Tungsten has been irradiated with 5 MeV α-particles and the production of vacancies at the surface observed. It is shown that this technique can be used to investigate the atomic structure at a grain boundary and around the core of a dislocation while it can be applied to the study of radiation damage to give information about the distribution and range of the damage produced by a single high-energy impact.

§ 1. Introduction

Müller (1956) has demonstrated that it is possible to resolve the atomic structure of a crystal lattice by means of the field ion microscope, but very little work has been published that makes use of this technique. Preliminary observations on oxidation, radiation damage, dislocations and vacancies have been made by Müller and Hudda (1960), and some results have been reported on adsorption (Ehrlich 1960). Müller and Good (1956) have proposed a theory of image formation and Müller (1958) and Drechsler and Wolf (1958) have discussed the interpretation of the micrographs. This interpretation involves two distinct stages. Firstly, the regions of high ion current which produce the bright spots in the image are related to the atomic projections on the specimen surface. Secondly, the atomic structure of the surface deduced from micrographs must be related to the atomic structure of the lattice in the bulk metal and in the absence of the high electric field. The first part of the analysis is based on the assumption that the contrast and resolution of the image of an atom on the screen are directly related to the bonding of the atom to the underlying lattice. Thus atoms in close-packed planes show little or no contrast, while atoms at the edges of these planes show good contrast. Atoms in a plane of relatively loose packing (e.g. {111} in the b.c.c. lattice) may be resolved with very good contrast while those on a plane of intermediate packing (e.g. {211} in b.c.c.) may be resolved with good contrast in one direction but only poorly resolved in another direction (see fig. 1†).

---

† On a grant from the Israel Atomic Energy Commission, Tel-Aviv, Israel.

The second part of the analysis, the relation of the surface structure of a specimen in the presence of an electric field to the bulk lattice, presents more difficulty. One basic problem is the stress generated by the electric field (Müller and Good 1956) which produces a hydrostatic expansion of the lattice. The forces acting on the tip are of the order of $F^2/8\pi$, where F is the field at the surface, and the stress may reach values of the order of $K/50$ where K is the bulk modulus. However, because the stress is largely hydrostatic it has little effect on dislocations in the lattice. A shear stress is necessary to produce dislocation movement, and though quite large shear stresses are present due to the tip being neither truly hemispherical nor elastically isotropic, they are usually insufficient to cause appreciable dislocation movement, at least in the absence of an oscillating field (Müller *et al.* 1959). The electric field produced at the surface is of the order of 5 volts per angstrom during normal operation. The potential difference across an atomic diameter produced by this field is of the same order of magnitude as the ionization potential of a free metal atom, and such high electric fields should modify the electronic structure at the surface. The field tends to push the conduction electron cloud farther into the metal, leaving an array of positive metal ions protruding from the matrix, to a first approximation creating an electrical double layer. The screening effect of the metal ions partly controls the penetration of the conduction electrons, but the resulting field distribution at the surface should also be a function of the Fermi surface in the metal as well as of the surface geometry. Hence it is to be expected that the quality of the picture would depend appreciably on the metal being studied. One result of this effect would be that for tips of equal size from different metals one would expect the picture to appear at different voltages; this in fact has been observed in tungsten and molybdenum. The present theory of image formation (Müller and Good 1956) takes no account of this effect of the electronic structure of the metal.

In the present paper the structure of lattice defects observed in the field ion microscope is discussed in the light of the experimental conditions under which the specimen is studied.

§ 2. EXPERIMENTAL

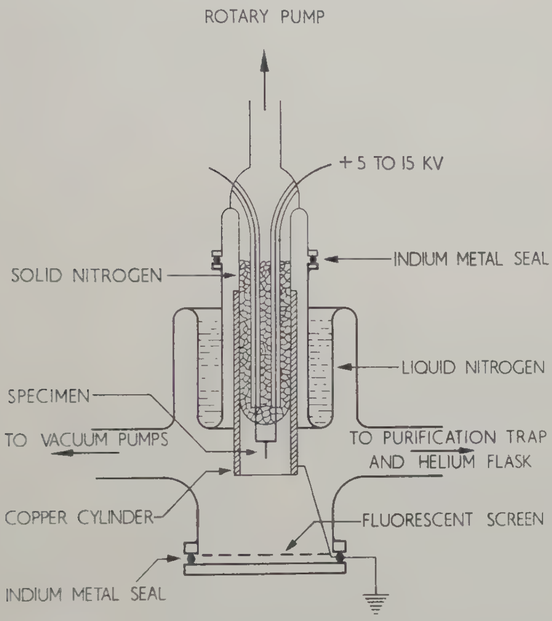
2.1. *Microscopy*

The apparatus is basically that used by Müller (1956) with the following modifications. As appreciable contamination of the specimen was found to occur at a background pressure of 10^{-6} torr the vacuum in the system was improved by winding asbestos-covered nichrome wire around the exposed glassware. After a 1 hr bakeout at 250°c the background pressure fell to about 5×10^{-8} torr. All grease seals were isolated from the microscope chamber by liquid nitrogen traps and the two seals actually in the microscope chamber were made with indium wire (fig. 2). The specimen support rods were cooled with liquid or solid nitrogen rather than with liquid hydrogen. Drechsler and Wolf (1958) have shown that adequate resolution can be obtained with liquid nitrogen provided the tip radius is limited to below 1000 Å. The helium was purified by passing through outgassed charcoal cooled by liquid nitrogen, and the gas pressure in the microscope kept at about 5×10^{-3} torr. The fluorescent screen was made demountable as ion bombardment damage in the phosphor eventually decreased the intensity of the image and resulted in prohibitively long exposure times for photography.

---

† Figures 1, 3, 4, 5, 8, 10 and 11 are shown as plates.

Fig. 2



Microscope chamber. (By courtesy of the Times Science Review.)

2.2. *Specimen Preparation*

Specimens of molybdenum, tungsten and platinum were prepared by electropolishing followed by field desorption in the microscope. The electropolishing techniques used were an extension of those applied to the production of thin metal films for transmission electron microscopy (Kelly and Nutting 1958). The polishing conditions for the different metals are given in the table. To obtain specimen tips of less than 500 Å radius without meniscus attack at the surface of the polishing bath it was necessary to lower the specimen wire into the solution during polishing. By controlling the rate of descent and the current density and by repeating the polishing operation a number of times, finely tapered tips were obtained. The specimens were examined at 2000 × magnification on a transmission bench microscope prior to mounting on the specimen holder of the field ion microscope. The criterion for a satisfactory tip was the disappearance of the Fresnel fringe at the tip of the specimen.

Metal	Electrolyte	Voltage	Temperature
Molybdenum	12·5% methanol 87·5% sulphuric acid	5 volts	20°c
Platinum	1·5% aqueous potassium cyanide	2–3 volts	20°c
	or molten sodium fluoride-potassium fluoride eutectic	3·5 volts	800°c
Tungsten	5% aqueous caustic soda	10 volts	20°c

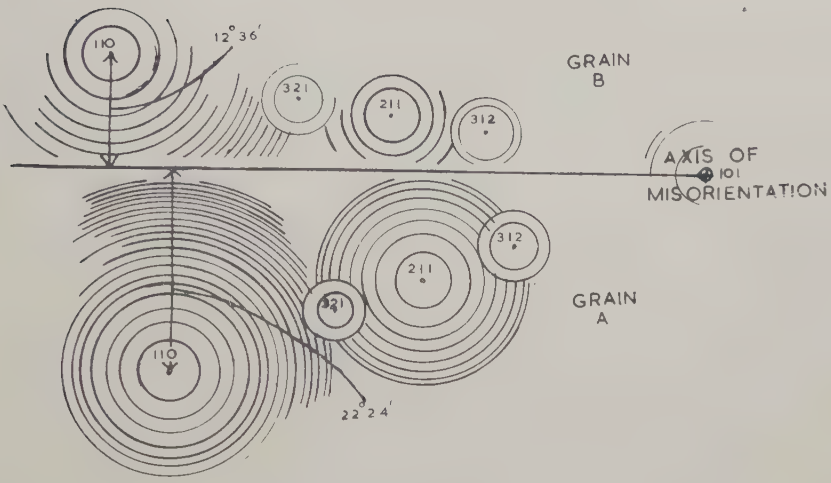
§ 3. Results

All pictures obtained from one metal showed certain characteristic features unique to the metal. Thus molybdenum never gave a picture of uniform intensity, but always showed a bright triangular region around the ⟨111⟩ poles (fig. 3). While tungsten, which produced a more uniform picture, showed bright bands running between the ⟨111⟩ poles (fig. 4). This is related to variations in surface energy, work function, polarizability and the energy of the Fermi surface, as well as to the radius of curvature.

Five distinct types of defect may exist in the crystal lattice. High angle grain boundaries, low angle grain boundaries (dislocation networks), dislocations, vacancies and dissolved impurity atoms. To these may be added interstitial atoms if radiation damage is being studied. In the present work several of these defects have been observed. They are (1) a high-angle grain boundary in molybdenum; (2) dislocations in molybdenum; (3) vacancies produced by α-particle bombardment of tungsten; (4) adsorbed impurity atoms tentatively identified as oxygen or nitrogen.

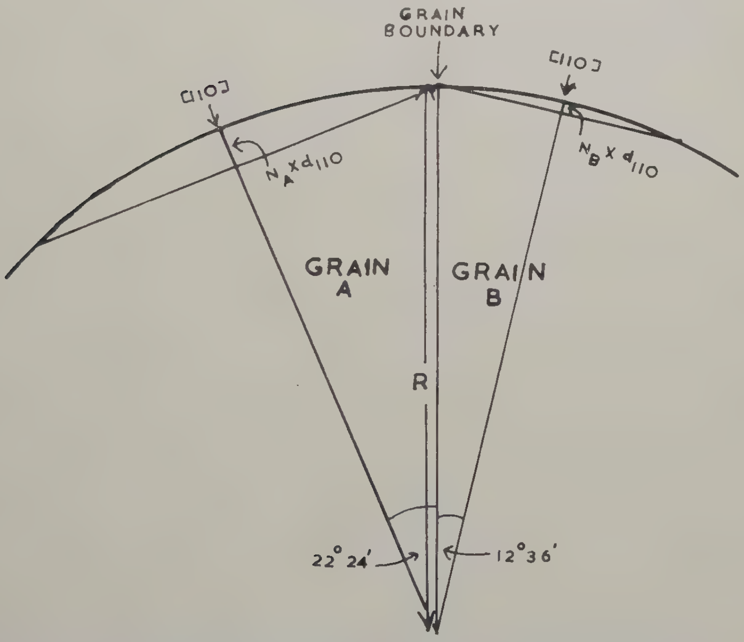
Interstitial atoms cannot be observed by this technique, since any interstitial atom will either assume a normal lattice position when brought to the surface or alternatively evaporate because of the high field gradient in its vicinity. It has not yet been established whether different atomic species can be identified, since the image contrast depends primarily on the extent to which an atom projects from the surface, and differences in atomic structure might be insufficient to distinguish between atoms of different species. An ordered alloy of known structure would enable an estimate to be made of the ability of the microscope to distinguish impurity atoms.

Fig. 6



Analysis of the grain boundary of fig. 5.

Fig. 7



Vertical section through the [110] poles across the grain boundary. R is calculated from $R(1-\cos\theta)=Nd$, where θ is the angle between the pole of a low index plane and any other known direction, N is the number of low index planes occurring in the angle θ and d is the spacing between the planes. The average value of R was found to be 560 Å.

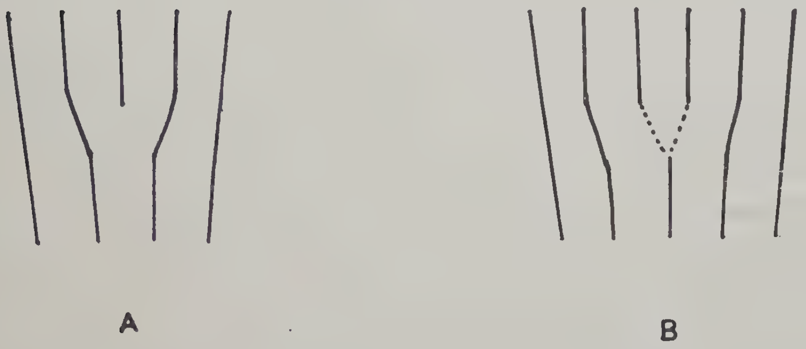
3.1. *High-angle Grain Boundary*

Because of the very small field of view ($\sim 10^{-10}$ cm^2) the chances of observing a grain boundary in the field ion microscope are very slight. One boundary has been observed in molybdenum, however (fig. 5). Figure 6 shows the crystallographic analysis of fig. 5, while a vertical section taken through the [110] poles in the two grains is shown in fig. 7. The axis of misorientation is about the [101] direction and the boundary is an almost symmetrical tilt boundary of angle 40°. The width of the boundary may be estimated from the spacing of the edges of the last complete (110) plane on the two sides of the boundary (figs. 5 and 6), it is seen to be about three lattice spacings, in accordance with theory (McClean 1957). The presence of large bright spots along the boundary may be taken to indicate segregation of impurities in this region, but there is no proof that this is not either impurity atoms adsorbed while the tip was under observation or else metal atoms giving exceptional contrast in a disordered zone.

3.2. *Dislocations*

Figure 8 (*a*) shows a dislocation in molybdenum. The dislocation is observed as an apparent missing half-plane of close-packed atoms, and there is no evidence of dissociation to form a stacking fault. The picture is from a region near the [110] pole and represents an oblique section through the close-packed (110) planes, each row of bright spots being the edge of a close-packed plane. The Burgers vector of the dislocation in fig. 8 (*a*) can be unambiguously identified from the missing half-plane, giving a displacement vector $\frac{1}{2}a[01\bar{1}]$, and the distortion of the planes immediately above and below the missing plane, giving a shear component displacement $\frac{1}{2}a[100]$. The actual Burgers vector is then $\frac{1}{2}a[100]+\frac{1}{2}a[0\bar{1}\bar{1}]=\frac{1}{2}a[1\bar{1}1]$. It is not possible to state whether the dislocation is predominantly edge or screw since the angle at which it comes to the surface is unknown, although this could be determined by careful desorption of successive (110) planes. In fact this is unimportant since it is only the edge component at the surface that can be accurately analysed. This is because on the atomic scale all dislocations are effectively normal to the free surface they intersect. Only lateral displacements can be accurately measured at the surface so only edge structure can be distinguished. This is not to say that a helical surface, due to the emergence of a screw dislocation, cannot be observed, but the core structure of a screw dislocation cannot be analysed in the field ion microscope. Figure 8 (*b*) shows the dislocation of fig. 8 (*a*) analysed on the assumption that it has been formed by the removal of half a (110) plane followed by a collapse and shear process of the type discussed by Kuhlman-Wilsdorf (1958). The lattice displacements of the close-packed planes can be measured fairly accurately and the core structure of a dislocation formed by the hypothetical collapse and shear process can be deduced. Figure 9 shows two possible core configurations in a plane normal to the (110)

Fig. 9



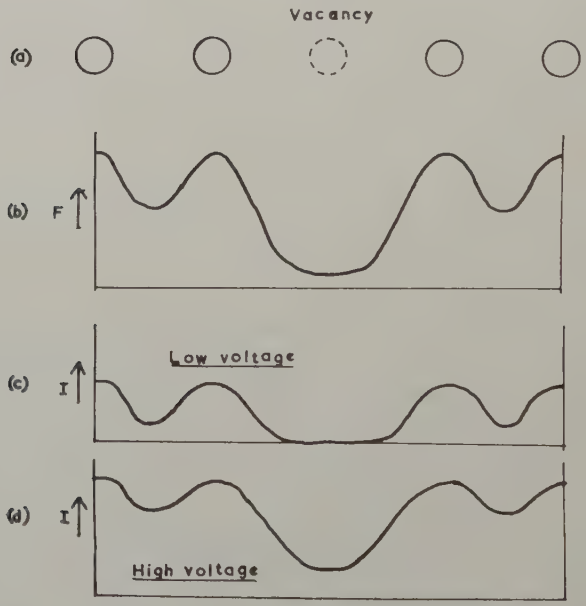
(A) Classical core configuration of an edge dislocation.
(B) Configuration deduced from the analysis of fig. 8 (*a*).

planes and containing the edge component of the Burgers vector $(\frac{1}{2}a[0\bar{1}\bar{1}])$. Figure 9 (A) is the configuration usually drawn in elementary textbooks. Figure 9 (B) is that deduced from fig. 8 (*a*) with the assumptions made above. Because of the atomic scale of the observations the actual configuration deduced should not be very sensitive to the orientation of the dislocation line with respect to the surface and it could be inferred that fig. 9 (B) is in fact the normal core structure in edge dislocations in molybdenum. The core structure should not be influenced by the hydrostatic stress due to the applied electric field since the elastic strain energy density due to the hydrostatic stress is only of the order of 10% of the strain energy density at the core of a dislocation.

3.3. *Vacancies*

Evidence for thermal vacancies in platinum quenched from just below the melting point has been obtained by Müller (1960). His observed concentration of 5×10^{-4} is of the right order of magnitude. However, the concentration of thermal vacancies in pure metals is always low and these studies are therefore not very rewarding. It was thought that vacancies might be studied in higher concentrations by bombarding the lattice with high energy α-particles. As explained previously, it is not possible to observe interstitial atoms. In molybdenum it was found impossible to prevent a small amount of desorption at normal working voltages, as the field strength required to obtain a good image is of the same order as the desorption field for this metal. In tungsten the desorption field is some 30% higher and a stable picture can be maintained for long periods of time. Thus, although accumulated damage was observed in molybdenum, specific defects attributable to a single impact could not be readily identified. A 0·5 millicurie polonium α-particle source was mounted about 1 cm from the tip. The 1 mm diameter source was positioned so that no particles could enter through the surface under observation, but it was possible for them to be scattered so as to emerge from this surface. At the surface focusing collision sequences (Silsbee 1957) result in the ejection of surface atoms (Nelson and Thompson 1961). If only one atom is ejected per replacement focusing sequence (Goland *et al.* 1960) there will be no change in the surface structure. However, as the energy at the start of a collision sequence is about 300 ev, more than one atom will be ejected unless the energy has fallen to below about 10 ev. This results in the production of vacant sites at the surface which mark the points of emergence of the collision sequences. Vacancies produced by knock-on collisions are less likely to be observed unless the specimen is desorbed after bombardment. If this is done vacancies present at the start of a replacement collision sequence will be revealed in addition to those produced by the knock-on mechanism. Figure 10 shows a tungsten surface before and after an impact. In fig. 10 (*a*) all the missing atoms in fig. 10 (*b*) are dotted while two atoms which have appeared in fig. 10 (*b*) are indicated similarly. The relative concentration of missing atoms in the region at the edges of the (110) planes suggests that the primary knock-on is underneath the surface in this area. 40 atoms have been displaced in the field of view in fig. 10. Taking the initial energy of a focusing collision as 300 ev and assuming that about 20% of the surface atoms are in fact

Fig. 12



Schematic behaviour of field strength and current density near a region containing a vacancy. (*b*) field distribution; (*c*) ion current density as a function of position at low voltages; (*d*) as (*c*) at a high voltage.

visible in the field ion image, the total energy dissipated in focusing sequences should be about 240 kev, which is approximately $\frac{1}{2}E$ max. and corresponds with the average energy imparted to a primary knock-on by an α-particle. Because the range of a primary knock-on is appreciable on the scale of fig. 10 there is no significant crystallographic distribution of the displacements. This would be expected to show maxima in the $\langle 111 \rangle$ and $\langle 100 \rangle$ directions measured from the centre of the disturbance, since these are the directions along which focusing is expected in a body-centred cubic lattice. Due to overlapping field ionization over neighbouring atoms vacancies show up with improved contrast when the voltage is increased (fig. 11). Figure 12 shows schematically how this effect arises from variations in the field strength over an atom, an interstitial hole and a vacancy. The surface vacancy concentration produced by sputtering (figs. 10, 11) is much greater than that observed after subsequent desorption. This is because there is a volume distribution of vacancies underneath the surface whereas the surface vacancies are at the points of emergence of focusing sequences.

3.4. *Adsorbed Atoms*

In analysing field ion micrographs it is necessary to distinguish adsorbed contamination from metal atoms in the original lattice. In some cases the bright atoms are readily identifiable as foreign atoms, for example the atom above the dislocation in fig. 8 (*a*). More often these brighter atoms are in positions where it is difficult to decide whether they are an artefact or not (e.g. fig. 10). In all cases where foreign atoms could be identified they were at the edges of low index planes where they had four or more nearest neighbours in the metal lattice.

§ 4. DISCUSSION

The main advantage of the field ion microscope is its outstanding resolution. In addition it has the advantage that the specimen can be observed and photographed during actual experiments. Moreover, radiation damage can be observed at a temperature where diffusion effects are negligible, so that thermally activated processes occurring after the initial thermal spike has dissipated itself are unimportant. However, because of the severity of the experimental conditions under which the microscope is operated, considerable care must be exercised before concluding that the observed structures are necessarily representative of the bulk lattice. It is usually safe to neglect the hydrostatic stress when observing lattice defects by this method as the strain energy density near a defect is appreciably greater than that introduced by the stress. This is true near the core of a dislocation and at grain boundaries (Read 1953). In these cases it is probably justifiable to extrapolate the surface observations to the bulk lattice. Thus it is mainly as a technique for observing the geometry and distribution of defects that cannot be resolved by more conventional techniques that field ion microscopy is useful.

Preliminary results have (*a*) confirmed the theoretical estimate of the width of a grain boundary; (*b*) revealed the core structure of dislocations; (*c*) provided evidence for the existence of focusing collisions and shown their distribution about a thermal spike. A consideration of the mechanism of image formation suggests that in each case considerable information may be derived from the micrographs and that it is possible to infer the distribution and structure of defects in the bulk lattice.

ACKNOWLEDGMENTS

This work was supported by the U.K.A.E.A. Our thanks are due to Professor A. H. Cottrell for the provision of laboratory facilities and for very helpful discussions. We should also like to thank M. J. Southon and Dr. M. Drechsler for useful discussions and M. J. Southon for assisting with part of the experimental work. We are grateful to the U.K.A.E.A. for permission to publish.

REFERENCES

DRECHSLER, M., and WOLF, P., 1958, *Proceedings of Fourth Int. Conf. on Electron Microscopy, Berlin* (Springer), p. 835.

EHRLICH, G., and HUDDA, F. G., 1960, *J. chem. Phys.*, **33,** 1253.

GOLAND, A. N., MILGRAM, M., VINEYARD, G. H., and GIBSON, J. B., 1960, *Bull. Amer. phys. Soc.*, **5,** 26.

KELLY, P. M., and NUTTING, J., 1958, *J. Inst. Met.*, **87,** 385.

KUHLMAN-WILSDORF, D., 1958, *Phil. Mag.*, **3,** 125.

MCCLEAN, D., 1957, *Grain Boundaries in Metals* (Oxford), p. 42.

MÜLLER, E. W., 1956, *J. appl. Phys.*, **27,** 474; 1958, *Proceedings of Fourth Int. Conf. on Electron Microscopy, Berlin* (Springer), p. 820; 1960, *Proceedings of Conf. on Structure and Properties of Thin Films* (New York: Wiley), p. 476.

MÜLLER, W. E., and GOOD, R. H., 1956, *Handb. Phys.*, **21,** 218.

MÜLLER, E. W., PIMBLEY, W. T., and MULSON, J. F., 1959, *Proc. of Symp. on Internal Stresses and Fatigue in Metals* (Elsevier).

NELSON, R. S., and THOMPSON, M. W., 1961, *Proc. roy. Soc.* A, **259,** 458.

READ, W. T., 1953, *Dislocations in Crystals* (McGraw-Hill), Ch. 13.

SILSBEE, R. H., 1957, *J. appl. Phys.*, **28,** 1246.

CORRESPONDENCE

The Observation of Four Types of Hall Constant Anisotropy in Copper and their Role in the Determination of the Fermi Surface

By J. E. Kunzler and J. R. Klauder
Bell Telephone Laboratories, Incorporated, Murray Hill, New Jersey

[Received April 19, 1961]

Four of five possible types of high field Hall constant anisotropy predicted for metals (Lifshitz and Peschanskii 1958) have been observed by us in copper and are reported in this letter. The fifth type is not expected for copper. The observed Hall constant behaviour is markedly different from the results reported for copper by earlier workers (Kittel 1956, Berlincourt 1958), primarily due to our use of samples that were both single crystals and highly pure ($R_{273^\circ\mathrm{K}}/R_{4\cdot 2^\circ\mathrm{K}}$ as large as 8000). Generally when the magnetoresistance saturates, the Hall constant can be directly related to the Fermi surface shape; the essential conditions are summarized and several Hall constant calculations, using the Pippard model, are compared with the corresponding experimental values. Finally, the utility of Hall voltage data for distinguishing between different possible Fermi surface topologies is illustrated for the first time.

The five types of Hall constant behaviour that result from theoretical considerations of the galvanomagnetic properties of metals in the high field limit are summarized in table 1. This classification is based in part on the electron- or hole-like nature of the carrier orbits and, in some cases, on the equality or the non-equality of the number of ‘ holes ’ and ‘ electrons ’. It is the fact that copper has an open Fermi surface (contacts the Brillouin zone boundary) that permits the existence of the first four types of behaviour and their observation in the appropriate magnetic field directions, rather than only the first type which would be characteristic of a free electron model for copper.

Type I behaviour is observed in copper when the magnetic field is well away from principal symmetry directions and is directed within those regions that correspond to the valleys of saturation on the stereograms of magnetoresistance data (Klauder and Kunzler 1960 a, b). Figure 1 shows the observed Hall voltage at 18 kgauss when the magnetic field was rotated in a plane perpendicular to the current axis of the sample, which was near [5$\bar{1}$1]. When $\theta = -3^\circ$, the magnetic field was perpendicular to a line connecting the contacts, and thus the Hall voltage decreases as cos $(\theta + 3^\circ)$. The envelope tends to follow the dotted curve which represents the theoretical behaviour that would be expected for one electron per atom if the Fermi surface did not contact the zone boundary. For $\theta > -40°$, the observed curve deviates from the dotted curve due to the effect of open orbits that are present each time the magnetic field crosses a magnetoresistance ridge.

Table 1. Five types of high field behaviour of the galvanomagnetic properties of metals and their relationship to 'electron' or 'hole' orbits

Type	Nature of Orbits	Behaviour of Resistivity in high magnetic field	Behaviour of Hall constant in high magnetic field
I	All closed†	$\rho \to$ saturate	$R_H \propto n^{-1}$
II	Negligible number of open orbits†	$\rho \to$ saturate	$R_H \propto (n_e - n_h)^{-1}$
III	Open orbits in only one direction	$\rho \propto H^2$‡	$R_H \to$ const‖
IV	Open orbits in more than one direction	$\rho \to$ saturate	$R_H \propto H^{-2}$
V	All closed $n_e = n_h$	$\rho \to H^2$§	$R_H \to$ const‖

n = number of carriers/atom; n_e = number of 'electrons'/atom; n_h = number of 'holes'/atom.

†$n_e \neq n_h$.

‡$\rho \to$ saturate if the current direction is perpendicular to the open orbit direction.

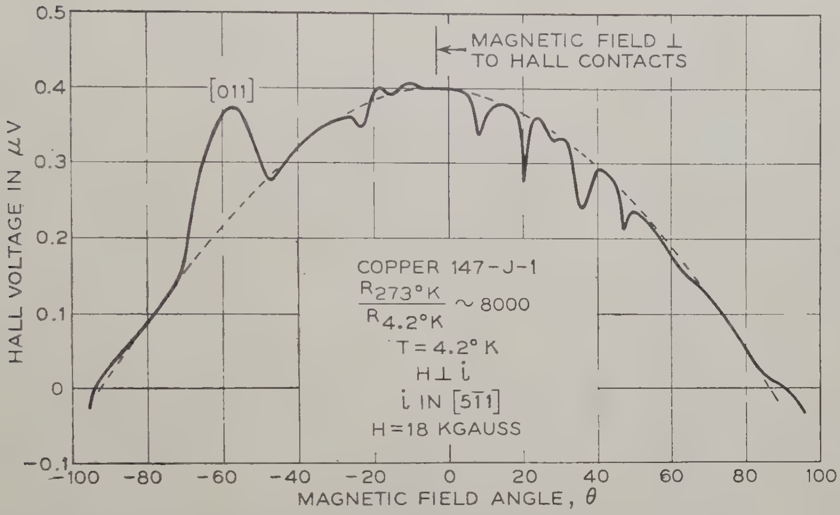
§$\rho \to$ saturate if the magnetic field is along the current direction.

‖R_H is not related to the number of carriers in any simple manner.

The smooth peak at $\theta = -55°$ (fig. 1) is a special case of Type II behaviour and represents the Hall constant when the magnetic field is along the [011] direction. In general, there is only a negligible number of open orbits possible when the magnetic field is along a principal symmetry direction. However, this is not quite the case for the ⟨110⟩ directions in copper. Even though there is a negligible number of open orbits present on either side of the [011] field direction at $\theta = -55°$ in fig. 1, which is typical of Type II behaviour, a small number of [100]-directed open orbits exist when the field is directly along the [011] axis, these open orbits being allowed by the twofold symmetry at this point. However, the smoothness of the observed Hall voltage peak in the [011] region plus the sharp dip characteristic of the magnetoresistance (Klauder and Kunzler 1960 a) permit us to ascribe the Hall constant at $\theta = -55°$ to the [011] symmetry axis itself. The value of the Hall constant derived from this observation, as well as values for other samples, are summarized in table 2.

Similar data for the ⟨100⟩ and ⟨111⟩ directions, which possess sufficiently high symmetry so that only a negligible number of open orbits are possible, were obtained from other curves, such as fig. 2. This figure represents the Hall voltage at 6 kgauss with the plane of rotation of the magnetic field transverse to the [110] current direction. In this figure the Hall constant can be derived for all three major crystal directions. We find that both the Hall constant and magnetoresistance are extremely sensitive to orientation when the magnetic field is near a ⟨111⟩ direction; this makes it very difficult to separate the part of the transverse voltage that is even in magnetic field from the Hall voltage, which is odd in field. Even when the magnetic field was reduced to 6 kgauss, as represented by fig. 2, the values of the Hall voltage in the ⟨111⟩ directions were somewhat uncertain and only an estimate can be made from the various data. Experimental refinements that are expected to overcome this difficulty are in progress.

Fig. 1



Observed Hall voltage as the magnetic field was rotated in a transverse plane. The long axis of the sample (current direction) is near the $[5\bar{1}1]$ direction. The solid curve is the recorded voltage with the contribution removed that is even in magnetic field. At $\theta = -3°$ the magnetic field direction is perpendicular to a line connecting the Hall contacts. The dotted curve is the theoretical voltage for one electron/atom ($R_{\mathrm{H}} = 7{\cdot}4 \times 10^{-12}$ volt cm/abamp gauss). The envelope represents Type I behaviour, the peak Type II, and the dips Type III as defined in table 1.

Type III behaviour is shown by each of the many depressions in the curve of fig. 1. As the magnetic field is rotated, such valleys appear in the Hall voltage each time a direction is crossed where open orbits exist (e.g. a magneto resistance ridge). All field directions between the symmetry axes of fig. 2 correspond to Type III behaviour, because the magnetic field is ‘ riding ’ along a magnetoresistance ridge. Thus,

Table 2. Observed and predicted values of the Hall constant for various magnetic field directions in copper. The value given is the Hall constant expressed in volt cm/abamp gauss and increased by a scale factor of 10^{12}

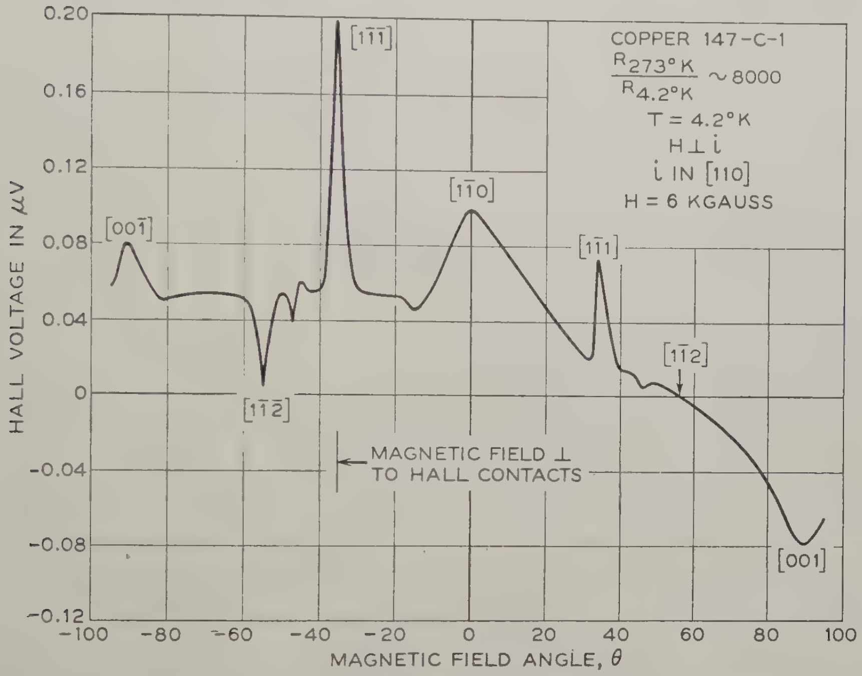
		Hall constant			'Electrons' / atom	
Direction of magnetic field	Sample	Observed		Predicted†	Observed	Predicted†
'Valleys'	147–J–1		7·4 ± 0·5	7·4	1·00	1·00
⟨100⟩	118–B–1	13·3				
	118–G–1	12·5	12·6 ± 0·5	10·3	0·59	0·71
	147–C–1	12·0				
⟨110⟩	118–C–1	11·7				
	147–C–1	11·3	11·8 ± 0·5	10·2	0·63	0·72
	147–J–1	12·3				
⟨111⟩	147–C–1		20 ± 5	22·7	0·37	0·32

† Based on Pippard's Fermi surface.

except at the symmetry directions, the changes in Hall voltage are gradual.

Type IV behaviour is apprent in fig. 2 when the magnetic field is along the $[1\bar{1}\bar{2}]$ direction ($\theta = -56°$). A similar dip is expected at $\theta = +56°$, but the Hall voltage goes to zero at this angle because the magnetic field is parallel to a line joining the Hall contacts. However, the dip at $\theta = +56°$ was observed on a second set of transverse contacts located 90° from those represented by fig. 2. The $[1\bar{1}\bar{2}]$-dip is a result of the simultaneous existence of open orbits in [110] and $[1\bar{1}1]$ directions when the magnetic field is along the $[1\bar{1}\bar{2}]$ axis.

Fig. 2



Observed Hall voltage as magnetic field was rotated in a transverse plane. The long axis of the sample (current direction) is the [110] direction. The curve is the recorded voltage with the contribution removed that is even in magnetic field. At $\theta = 0°$ the magnetic field is along the $[1\bar{1}0]$ direction. At $\theta = 56°$ the magnetic field is along a line joining the Hall contacts. The Hall constant near the ⟨111⟩ directions is so sensitive to orientation that it is difficult to obtain accurate data. The $[1\bar{1}\bar{2}]$ direction shows Type IV behaviour, the other symmetry directions that are labelled represent Type II behaviour, and the regions between the principal directions correspond to Type III behaviour.

The Hall constant for Types I and II behaviour is a measure of the difference in the number of carriers having electron-like properties and the number having hole-like properties. In these two cases the Hall constant R_{H} is related to the Fermi surface shape by

$$R_{\mathrm{H}}=\frac{h^3}{2ec}\left\{\int dp_z[S_e(E_f,p_z)-S_h(E_f,p_z)]\right\}^{-1}, \quad . \quad . \quad . \quad (1)$$

where the z direction is chosen parallel to the magnetic field. S_e and S_h are closed areas formed by the intersection with the Fermi surface of a plane perpendicular to p_z. Electron-like properties are associated with areas enclosing states having energies less than the Fermi energy and hole-like properties with areas enclosing states having energies greater than E_f. It is apparent that the Hall constant depends on the detailed shape of the Fermi surface and that values for several non-equivalent directions place severe constraints on the shapes of possible surfaces. Numerical calculations of the Hall constant for several orientations of copper have been made (Klauder 1961) using the Pippard–Fermi surface shape (Pippard 1957, Garcia-Moliner 1958) and are tabulated in table 2. The agreement with experiment is satisfactory, but can be improved by making the ‘ arms ’ or ‘ necks ’ somewhat thicker.

The important role that Hall voltage data can play in deciding among possible Fermi surface topologies is well illustrated by the case of copper. A topological model for the Fermi surface of copper that has arms in the $\langle 100\rangle$, $\langle 110\rangle$ and $\langle 111\rangle$ directions was proposed (Alekseevskii and Gaidukov 1959) on the basis of magnetoresistance data. The gross features of our data (Kunzler and Klauder 1960; Kunzler *et al.* 1960) are compatible with their model, and by making use of the concept of higher order open orbits (Klauder and Kunzler 1961) our data is also compatible with the Pippard model. However, the ‘ fine structure ’ found in our data favours the Pippard model. Finally, it is the Hall voltage dip in the $[1\bar{1}2]$ direction of fig. 2 that rules out the Russian model because arms in both the $\langle 110\rangle$ and $\langle 111\rangle$ directions would not have permitted the simultaneous existence of open orbits in more than one direction as is evidently essential when the magnetic field is along any $\langle 112\rangle$ direction. It is apparent that Hall data, especially when it is used in conjunction with magnetoresistance data, can play an important role in the unravelling of the Fermi surfaces of the more complicated metals.

REFERENCES

ALEKSEEVSKII, N., and GAIDUKOV, YU, 1959, *J. exp. theor. Phys.*, **36,** 447; 1959, *Soviet Phys. JETP*, **9,** 311.

BERLINCOURT, T., 1958, *Phys. Rev.*, **112,** 381.

GARCIA-MOLINER, F., 1958, *Phil. Mag.*, **3,** 207.

KITTEL, C., 1956, *Introduction to Solid State Physics*, Second edition (New York: John Wiley & Sons, Inc.), p. 298.

KLAUDER, J. R., 1961, *Bell Syst. tech. J.*, Sept. 1961.

KLAUDER, J. R., and KUNZLER, J. E., 1960 a, *Proceedings of the Fermi Surface Conference* (New York: John Wiley & Sons, Inc.); 1960 b, *Proceedings of the VIIth International Low Temperature Physics Conference, Toronto, Canada*; 1961, *J. Phys. Chem. Solids*, **18,** 256.

KUNZLER, J. E., and KLAUDER, J. R., 1960, *Bull. Amer. phys. Soc.*, **5**, 150.
KUNZLER, J. E., KLAUDER, J. R., and BRENNERT, G. F., 1960, *Proceedings of the VIIth International Low Temperature Physics Conference, Toronto, Canada.*
LIFSHITZ, I., and PESCHANSKII, V., 1958, *J. exp. theor. Phys.*, **35**, 1251; 1959, *Soviet Phys. JETP*, **8**, 875.
PIPPARD, A. B., 1957, *Trans. roy. Soc.* A, **250**, 325.

Surface Detachment from Polyethylene Crystals

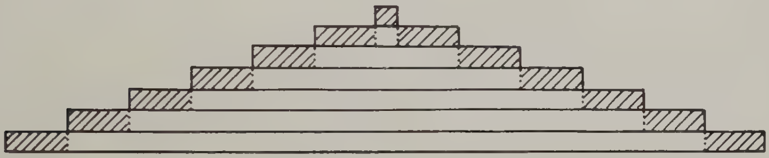
By D. C. Bassett
H. H. Wills Physics Laboratory, University of Bristol, Bristol 8, England

[Received March 30, 1961]

§ 1. Introduction

During recent work on polyethylene crystals (Keller and Bassett 1960, Bassett and Keller 1961 a, b) a method has been discovered for detaching from a multilayer crystal those parts of layers forming the external fold surface (the surface layer, see fig. 1†). Thus is done by 'fixing' the crystal surface by the evaporation on it of a carbon layer and then dissolving away the unfixed part of the crystal under controlled conditions to give what we propose to call a detachment replica. In a more advanced form, the technique permits the dissolution of only the thinner layers of crystals leaving thicker ones attached to the detachment replica. These effects alone provide a useful new tool for investigating crystalline polymers. In addition, however, the crystal layers of the detachment replica diffract electrons as do polyethylene single crystals. Thus it is now possible to remove the exposed surface layer of a thick specimen and to conduct diffraction studies upon it revealing orientation information which has so far been inaccessible to any other technique (Keller and Bassett 1960).

Fig. 1



Schematic representation of the 'surface layer' (shaded) which is removed from a multilayer polyethylene crystal in a detachment replica.

§ 2. Experimental

The technique to be described was discovered during work on polyethylene crystals having a discontinuous change of layer thickness (Keller and Bassett 1960). Their preparation is a little involved.

---

† Figures 2–6 are shown as plates.

Initially, truncated lozenge monolayer crystals, 150 Å thick, were grown at 90°C from an 0·01% solution in xylene of the linear polyethylene Marlex 50. Because the polymer is still partially soluble in xylene in this region of temperature, the crystals have to be filtered from suspension at 90°C to avoid the secondary crystallization which occurs on cooling (Bassett and Keller 1961 b). After filtration just to a 'gel-state' the crystals are removed from the filter and shaken up with cold, fresh xylene, forming a new crystal suspension (Bassett and Keller 1961 b). On heating this suspension the crystals are partially soluble below about 97°C where they dissolve completely. As is described in detail elsewhere (Keller and Bassett 1960, Bassett and Keller 1961 b), on heating the suspensions to, say, 93 or 94°C, the edges of the truncated lozenges dissolve and this dissolved material can be re-precipitated on cooling. It does so both as a thinner border to the original lozenges and as new crystals of the same thickness as the border. This is the growth history of the crystals shown in figs. 2, 3 and 4 (the new crystals have a dendritic habit). The original truncated lozenges form a characteristic fold in their interior when dried down, where the crystal is three layers thick (Bassett *et al.* 1959, Niegisch 1959). This has been interpreted as the result of the collapse of a hollow pyramidal crystal and is present in the composite crystals under discussion.

The specimen preparation is as follows: a drop of the crystal suspension is dried down on a glass or mica surface, a layer of carbon and then, if desired, metallic shadow is evaporated on it, and the carbon film plus crystals floated off on a water surface to be picked up on electron microscope grids. Various detachment replication effects can be obtained depending on how the crystals are dissolved away. For those as in fig. 3 the carbon film is held in contact with xylene at 98°C to, say, 105°C for about 15 sec, whereas 15 sec at, say, 90°C to 95°C gives those as in fig. 4. The temperature limits quoted are strict at 95°C and 98°C which straddle the dissolution temperature of the 150 Å thick layers, but less so for the other two. At temperatures above about 105°C one begins to get true replicas and below 85° or 90°C (depending on the layer thickness) no detachment replication at all.

All these detachment replicas give good crystalline diffraction patterns under appropriate viewing conditions (Agar *et al.* 1959). Figure 6 is a diffraction pattern from a detachment replica similar to that in fig. 5, showing growth on cooling around a truncated lozenge seed crystal. The slight misorientation of overlying layers seen visually is revealed. Attempts to obtain detachment replicas with crystals which have been preshadowed (Au : Pd, 40 : 60) before carbon was evaporated on them, are only partially successful, as the surface layers are blotchy. Moreover, their diffraction patterns are more or less disoriented, especially above 97°C where the crystals would ordinarily dissolve, and bear little or no relation to the original crystal outlines, i.e. recrystallization has occurred. Most of our work has, therefore, been confined to the specimen preparation outlined earlier.

§ 3. Discussion

The simplest type of detachment replica is that of fig. 3. Here there are only two contrast tones (which should be compared with a normal transmission electronmicrograph as in fig. 2). In certain places it is clear that these correspond to the thicknesses of interior and border (150 and 110 Å respectively). However, there are features where the crystal was obviously not originally a monolayer which also give only monolayer contrast. One is forced to the conclusion that everywhere only one layer remains. The electronmicrographs are consistent with this being the uppermost or surface layer, i.e. that in contact with the carbon film. Thus, where a thin layer is above a thick one, the contrast will be that for a thin layer and vice versa. The contrast in fig. 3 is completely explicable in this way.

It is obvious that the evaporated carbon film has given considerable extra stability against dissolution to the crystal layers with which it is in contact. In fact, it has raised their dissolution temperature by 10°c or so for the 150 Å layers and more for the thinner ones which ordinarily dissolve at lower temperatures than the thick ones (Lauritzen and Hoffman 1960, Bassett and Keller 1961 b). It is not known to what this extra stability is due. One may suggest that free radicals in the evaporated carbon film have linked it to the fold surface. This linking might be to the chain folds themselves, in analogy with recent radiation experiments (Salovey and Keller 1961) or possibly to chain-ends present on the fold-surfaces (Frank and Tosi 1961). For present purposes, however, it is sufficient merely to regard the carbon film as 'fixing' the crystal surface. The effects found with preshadowed crystals do not necessarily imply that metal shadow is 'fixing' the crystal surface. The observations seem at least qualitatively compatible with the carbon film 'fixing' the surface between shadow granulations. This would allow recrystallization of the unfixed part of the surface layers to occur in the way found.

Figure 4 shows effects obtained by dissolution at temperatures where the 150 Å layers are only partially soluble. These can be interpreted as due to a surface layer, plus all 150 Å layers which, for the short dissolution time, are effectively insoluble. Thus, where a new dendritic crystal overlays the interior, one sees the contrast due to this plus the thicker underlying layer, but outside where it overlays the border, only that for one layer, the thin underlying one having dissolved away. This explanation accounts for all other features such as the strong contrast of the internal central folds and, rather prettily, for the change of contrast of the growth pyramid on passing through the interior layer.

Selective dissolution can be extended to dissolve away layers of less than a certain thickness provided dissolution is sufficiently carefully controlled. It has been possible, for example, to dissolve away 110 Å layers leaving ones of 125 Å intact (these were ones grown at 70 and 80°c respectively). In this way, the technique provides a valuable means of distinguishing growth stages from each other.

Its greatest value, however, would seem to lie in applications such as that of fig. 5. This shows a crystal formed round a truncated lozenge seed which has been left undissolved, as is very clearly revealed by the contrast differences. It is found systematically in this preparation that the new growth adopts the orientation of the seed crystal. As the seed itself is still present in the detachment replica below the surface layer, it must be attached to it in some way. In fig. 5 part of the seed (the two re-entrant corners) is in contact with the carbon film itself, but in other cases there is no such direct contact. In this way, one may show up interlamellar connections which are, for example, of importance in the understanding of complex solution grown crystals and their relation to melt-crystallized material (Salovey and Keller 1961).

The detachment replica, moreover, can still diffract electrons, and as it is (except for the seed) a monolayer all over, it does so with much less bias against the small layers of the growth pyramids than there would be in a normal transmission diffraction study. (Figure 6 is an electron diffraction pattern from a detachment replica similar to that of fig. 5.) In fact, a large part of the original crystal was over 1000 Å thick and would have been too thick for transmission electron diffraction, especially since the crystals are beam sensitive (Agar *et al.* 1959). We are thus entering a region where orientation information has previously been unobtainable (Keller and Bassett 1960). For even thicker and more complex crystals (e.g. those in fig. 15 of Agar *et al.* 1959) this technique has already been used successfully for diffraction work which was not possible previously.

Finally it is conceivable that the technique could be extended to the lamellae in melt-crystallized polymers (Fischer 1957). In this way one might hope to establish the presence of a folded-chain configuration in a similar manner to that for solution-crystallized specimens (Keller 1957). The possibility is currently being investigated in this laboratory.

ACKNOWLEDGMENTS

I wish to express my thanks to Professor F. C. Frank, F.R.S., and to Dr. A. Keller for their help and encouragement throughout this work. I wish also to acknowledge the award of a Postgraduate Scholarship by the Trustees of the Courtaulds Scientific and Educational Trust Fund.

REFERENCES

AGAR, A. W., FRANK, F. C., and KELLER, A., 1959, *Phil. Mag.*, **4,** 32.
BASSETT, D. C., FRANK, F. C., and KELLER, A., 1959, *Nature, Lond.*, **184,** 810.
BASSETT, D. C., and KELLER, A., 1961 a, *Phil. Mag.*, **6,** 345 ; 1961 b (to be published).
FISCHER, E. W., 1957, *Z. Naturf.* A, **12,** 753.
FRANK, F. C., and TOSI, M. P., 1961, *Proc. roy. Soc.* (in the press).
KELLER, A., 1957, *Phil. Mag.*, **2,** 1171.
KELLER, A., and BASSETT, D. C., 1960, *J. R. micr. Soc.*, **79,** 243.
LAURITZEN, J. I., and HOFFMAN, J. D., 1960, *J. Res. nat. Bur. Stand.* A, **64,** 73.
NIEGISCH, W. D., 1959, *J. Polym. Sci.*, **40,** 263.
SALOVEY, R., and KELLER, A., 1961, *Bell Syst. tech. J.*

Antinucleation Processes in Ferromagnetics

By A. Jaśkiewicz
Low Temperature Laboratory, Wrocław University

[Received February 19, 1961]

§ 1. Introduction

The nucleation mechanism of new domains leading to a change in domain orientation was proposed by Becker and Döring (1939). This mechanism was developed further by Kondorsky (1949) and Goodenough (1954). Goodenough considered in great detail the influence of granular inclusions, grain boundaries and lamellar precipitates on the nucleation of domains of reverse magnetization. The nucleation process was facilitated by the magnetic pole distribution around the imperfections. However, in the nucleation processes treated so far, new domains are thought of as being created in the existing domain structure, independently of the volume and shape associated with this structure.

In the present paper a magnetization reversal mechanism is presented which takes into account the volume and shape of the existing domains. The externally applied magnetic field causes some ferromagnetic domains to be destroyed, transforming them to a temporary superparamagnetic state, subsequently new domains are formed which are oriented in a favourable direction.

§ 2. The Free Energy of a Domain

The free energy of a ferromagnetic domain may be written in the general form

$$F=F_v+F_w+F_d+F_a+F_i, \qquad (1)$$

where F_v is the volume energy, F_w the surface energy, F_d the demagnetization energy, F_a the anisotropy energy and F_i the field interaction energy (for further details see Kittel and Galt 1956).

The volume energy may be taken to be the difference between the free energies of the paramagnetic and ferromagnetic phases, the energy density being given by a general expression of the form $A(I_s, T)$, where I_s is the spontaneous magnetization and T the temperature. The surface energy density is denoted by σ. The demagnetization energy depends on the shape of the domain; for a domain having spheroidal shape, with $u=l/r$, where r is the minor axis, l the major axis, the demagnetization energy is given by

$$F_d=\frac{32}{3}\pi^2 J_s{}^2\eta r^3 u f(u), \qquad (2)$$

where

$$f(u)=\frac{1}{u^2-1}\left[\frac{1}{2}\frac{u}{\sqrt{(u^2-1)}}\ln\frac{u+\sqrt{(u^2-1)}}{u-\sqrt{(u^2-1)}}-1\right]$$

and η a coefficient depending on the magnetic permittivity.

Assuming for simplicity that the magnetization vector of the domain is oriented along the easy magnetization axis and that the direction of an external magnetic field is also oriented along this axis, the expression for the free energy may be written in the form

$$F=\frac{4\pi}{3}ur^3A+\pi^2ur^2\sigma+\frac{32}{3}\pi^2J_{\mathrm{s}}^2\eta r^3uf(u)\pm\frac{4\pi}{3}r^3uJ_{\mathrm{s}}H, \qquad (3)$$

where the + and − signs are for domains with magnetization antiparallel and parallel to the external magnetic field.

§ 3. CRITICAL DIMENSIONS OF A DOMAIN

From nucleation theory (see, for example, Turnbull 1956) it is known that the nucleus of a new phase is unstable and vanishes when it is smaller than a critical size. Only nuclei larger than this size are stable and can grow. For a domain of a given shape parameter *u* the critical dimensions can be calculated as a function of the external field, leading to the criterion whether a domain of a given volume and shape can grow or disappear.

From (3) the critical dimensions for a domain with shape parameter *u* and with parallel (p) and antiparallel (a) magnetization referred to an externally applied magnetic field are

$$r_{\mathrm{c}}^{(\mathrm{p})}=-\frac{\pi\sigma}{2A+16\pi J_{\mathrm{s}}^2\eta f(u)-2J_{\mathrm{s}}H}, \quad . \quad . \quad . \quad . \quad (4)$$

$$r_{\mathrm{c}}^{(\mathrm{a})}=-\frac{\pi\sigma}{2A+16\pi J_{\mathrm{s}}^2\eta f(u)+2J\ H}. \quad . \quad . \quad . \quad . \quad (5)$$

The negative value of volume energy density $A(I_{\mathrm{s}}, T)$ makes r_{c} positive.

From expressions (4) and (5) it appears that the critical dimensions are changed by the field. The antiparallel domain, though stable in zero field, may become unstable when the applied field is high enough. This is not caused by a change in the critical dimensions. Until the field reaches a critical magnitude, H_{c} say, corresponding to this critical dimension, the domain remains unaffected. It changes, however, when the field reduces the size to less than critical dimensions. This process may be called antinucleation.

§ 4. CONCLUSIONS

The above process does not exclude another elementary magnetization reversal process, namely wall displacement. The wall movement may facilitate or hinder the antinucleation or nucleation by changes of the shape and volume of a particular domain. It is obvious that the changes in volume caused by the wall movement will manifest themselves in magnetization changes.

The antinucleation process provides a new physical picture of magnetization reversal in ferromagnetics. The most characteristic feature is that account is taken of the domain structure which is present before the magnetic field is applied. The second feature is that this process also occurs in single crystals without imperfections.

The nucleation process which takes place when antiparallel domains are destroyed is promoted considerably (Jaskiewicz 1957) because the new domains nucleate in the superparamagnetic phase.

In later papers this mechanism will be described in more detail. It will be shown that Barkhausen discontinuities, superparamagnetism (Bean 1955) and other phenomena connected with the changes in domain structure, may be well understood in terms of the above picture.

ACKNOWLEDGMENTS

The author is greatly indebted to Professor J. Mazur, Head of the Low Temperature Laboratory, for valuable advice and help throughout. He also wishes to express his hearty thanks to Professor J. Nikliborc and Professor B. Makiej for their steady interest.

REFERENCES

BEAN, C. P., 1955, *J. appl. Phys.*, **26,** 1381.
BECKER, R., and DÖRING, W., 1939, *Ferromagnetismus* (Berlin: Springer).
GOODENOUGH, J. B., 1954, *Phys. Rev.*, **95,** 917.
JAŚKIEWICZ, A., 1957, *Acta phys. polon.*, **16,** 169.
KITTEL, C., and GALT, J. K., 1956, *Solid State Phys.*, **3,** 437.
KONDORSKY, E., 1949, *C.R. Acad. Sci. U.R.S.S.*, **68,** 37.
TURNBULL, D., 1956, *Solid State Phys.*, **3,** 226.

Dimer Problem in Statistical Mechanics — An Exact Result

By H. N. V. Temperley
Atomic Weapons Research Establishment, Aldermaston, Berkshire

and Michael E. Fisher
Wheatstone Physics Laboratory, King's College, London

[Received August 10, 1961]

An important, even though physically oversimplified model of a system (e.g. solution or gas) containing diatomic molecules is that of a lattice occupied by ‘ rigid dimers ’ (e.g. Fowler and Rushbrooke 1937). Each dimer fills two nearest-neighbour sites on the lattice and no site may be occupied by more than one dimer. This represents a statistical problem of some difficulty which has so far been solved exactly only in one dimension (Green and Leipnik 1960, Fisher and Temperley 1960). (It has been treated by various authors on the basis of approximate statistical methods, e.g. Rushbrooke *et al.* (1953).) This note reports an exact solution for the dimer problem on the plane square lattice in the limiting case where the dimers completely fill the lattice (close-packed or high density limit).

Consider the distribution of dimers on a plane square lattice of m rows and n columns. If $g_{mn}(l, k)$ is the number of ways of placing l horizontal or X-dimers and k vertical or Y-dimers on the lattice then the required configurational grand partition function is

$$Z_{mn}(x, y) = \sum g_{mn}(l, k) x^l y^k \qquad (1)$$

where x and y are the activities of X-dimers and Y-dimers respectively. The numbers of dimers l and k are subject to $l+k=\frac{1}{2}mn$.

For a large lattice

$$Z_{mn}(x, y) \sim [Z(x, y)]^{mn} \qquad (2)$$

and the free energy per lattice site approaches a limit

$$-(F/kT) = \lim\,(mn)^{-1} \ln Z_{mn}(x, y) = \ln Z(x, y). \qquad (3)$$

This is found to be

$$\ln Z(x, y) = \tfrac{1}{4} \ln 2 + (2\pi)^{-2} \int_0^\pi \int_0^\pi \ln\,(x^2 + y^2 - x^2 \cos\alpha - y^2 \cos\beta)\, d\alpha\, d\beta. \qquad (4)$$

The integral in this result is closely related to Onsager's (1944) expression for the partition function of the asymmetric Ising model *at* the critical temperature T_c. From (4) the free energy and entropy are seen to be continuous smooth functions of the activities x and y and hence of the relative density of X-dimers and Y-dimers. In the symmetric case

$x=y=1$ the integral simplifies and the number of arrangements of N dimers is asymptotically,

$$[Z(1,\,1)]^{2N}=[\exp(2G/\pi)]^N=(1{\cdot}791623)^N \qquad . \quad . \quad . \quad (5)$$

where $G=1-3^{-2}+5^{-2}-7^{-2}+\ldots=0{\cdot}915\,966$ is Catalan's constant. (The Bethe approximation (Chang 1939) yields the estimate

$$[Z(1,\,1)]^2\simeq 27/16=1{\cdot}6875.$$

Fowler and Rushbrooke's estimate (1937) was 1·80 which is very close to the exact result.)

The calculation leading to (4) is based on an algebraic property of the Pfaffian. This is a triangular array of numbers filling the spaces above the main diagonal in a determinant. It is evaluated, like a determinant, in terms of the elements of its top row and their corresponding minors, with the same sign conventions. The minor of a_{pq} is now defined by striking out from the Pfaffian both the pth and qth rows *and* columns. This means that each term in the expansion of a Pfaffian of order $2N$ will contain just N factors, and each index p will appear once and once only in such a product. Thus, if we number our $2N$ lattice points in any manner, and put $a_{pq}=x$ when points p and q are horizontal neighbours and $a_{pq}=y$ when p and q are vertical neighbours, all other a's being put equal to zero, it is evident that there will be a 1–1 correspondence between the terms of the Pfaffian and the permissible arrangements of dimers on the lattice of $2N$ points.

As usual in these problems we have to ensure that all the terms are positive. In particular, for the plane square lattice we can take the x's all positive and the y's of alternating signs *or* we can number alternate rows of the lattice in opposite senses. (So far, we have not solved the very much more interesting problem of a lattice with some points empty —the formal 1–1 correspondence can be maintained, but the signs of some of the terms go wrong.)

The Pfaffian has various properties that make it important in mathematical physics. It is, for example, a compact expression for the trace of a product of anticommuting ‘ Dirac operators ’ and its square is the antisymmetric determinant obtained by reflecting the Pfaffian in the main diagonal (Hurst and Green 1960, Caianiello and Fubini 1952). In the present problem, this determinant is almost cyclic and thus can be evaluated, even for a finite lattice with edges, by standard methods, the results closely resembling Kaufman's expressions (1949) for the finite Ising model. By way of example the number of ways of covering an 8 × 8 chessboard with dominoes has been found to be

$$12\,988\,816=16(901)^2.$$

In the symmetric case $x=y$, the asymptotic expansion is

$$\ln Z_{mn}(1,\,1)=mn(G/\pi)-(m+n)[\tfrac{1}{2}\ln(1+2^{1/2})-G/\pi]+C+O(m^{-1},\,n^{-1}) \qquad . \quad . \quad . \quad (6)$$

where C is a constant. The first term in the expression represents the 'bulk' free energy, whilst the second term is the 'boundary' free energy and would have a different value if the lattice had been wrapped on a torus. (This possibility can also be dealt with.)

The principal results of this note were discovered independently by the two authors. The detailed derivations are being prepared for publication.

Since completing this work we have heard (private communication) that P. W. Kasteleijn has obtained similar results.

ACKNOWLEDGMENTS

One of us (M. E. F.) would like to thank Dr. C. A. Hurst for a stimulating discussion, while the receipt of a preprint of Hurst and Green's paper stimulated H. N. V. T.'s interest in the problem.

REFERENCES

CAIANIELLO, E. R., and FUBINI, S., 1952, *Nuovo Cim.*, **9,** 1218.
CHANG, T. S., 1939, *Proc. roy. Soc.* A, **169,** 512.
FISHER, M. E., and TEMPERLEY, H. N. V., 1960, *Rev. mod. Phys.*, **32,** 1029.
FOWLER, R. H., and RUSHBROOKE, G. S., 1937, *Trans. Faraday Soc.*, **33,** 1272.
GREEN, H. S., and LEIPNIK, R., 1960, *Rev. mod. Phys.*, **32,** 129.
HURST, C. A., and GREEN, H. S., 1960, *J. chem. Phys.*, **33,** 1059.
KAUFMAN, B., 1949, *Phys. Rev.*, **76,** 1244.
ONSAGER, L., 1944, *Phys. Rev.*, **65,** 117.
RUSHBROOKE, G. S., SCOINS, H. I., and WAKEFIELD, A. J., 1953, *Faraday Soc. Disc.*, No. 15, 57.

A Note on Abnormalities in the Travel Time of a Wave between two Extensive Apertures

By G. Bradfield and E. T. Goodwin
National Physical Laboratory, Teddington

[Received May 3, 1961]

It is a common practice to attempt to measure the velocity of mechanical waves in a solid by putting two circular coaxial plane electromechanical transducers on two parallel faces distance R apart and measuring the time for transmitting a sudden disturbance to travel from one to the other, the assumption being made that the velocity c is R/t. Apart from other difficulties which are encountered in such an experiment there is a fundamental source of error due to a peculiar integration effect when each transducer, as is usually the case, is very many wavelengths across and a train of sinusoidal waves is employed : the apparent time of travel of the wave when plotted against distance is not a smooth curve but has steps in it. It will be appreciated that the arrival of the wave as a whole must be delayed compared with the time R/c, because some of the energy must proceed by a path longer than R from the tip of one transducer obliquely across to a point on the other transducer and this results in an angle of lag of the received wave which is a function of the geometry of the system, the wavelength and the transducer diameter or aperture $2a$.

Making the assumption of uniformity of excitation over the initial aperture and confining attention to a pair of transducers or apertures of the same size, we find that, provided the aperture of the radiator is greater than about 20 wavelengths ($20\,\lambda$), we can compute a relation between the angle of lag of the radiation and the range measured in units of $S = a^2/\lambda$ which is substantially independent of this aperture ; this angle of lag will now be tabulated against range to illustrate our point. The broad theory used has been discussed by Huntingdon *et al.* (1945) and by Seki *et al.* (1956) and others, but none of the workers in this field have noticed the consequences of the theory which are described in the present note, though a similar effect in the amplitude of the pressure has been investigated (Seki *et al.* 1956).

In addition to a general diminution of the angle of lag with range, the relation ultimately becoming hyperbolic, it will be observed that this

Relation of angle of lag to range R/S for $2\pi a/\lambda = 200$

Range R/S	Angle of lag (degrees)	Range	Lag angle maxima and minima (degrees)	
0·0	90·000			
0·2	83·591			
0·3	81·986			
0·4	80·779	0·448	79·772	(Min)
0·5	79·589	0·469	79·916	(Max)
0·6	78·242	0·581	78·206	(Min)
0·7	77·087	0·616	78·267	(Max)
0·8	75·486	0·826	75·434	(Min)
0·9	75·525	0·894	75·527	(Max)
1·0	74·759			
1·1	72·920			
1·2	71·084			
1·4	69·464	1·443	69·432	(Min)
1·6	69·588	1·640	69·604	(Max)
1·8	69·250			
2·0	67·751			
2·2	65·383			
2·4	62·573			
2·6	59·624			
2·8	56·711			
3·0	53·922			
3·2	51·298			
3·4	48·852			
3·6	46·584			
3·8	44·484			
4·0	42·542			
5·0	34·768			
6·0	29·295			
7·0	25·274	At longer ranges than this the product of R/S and angle of lag appears to be a constant within an accuracy of 1%.		
8·0	22·206			
9·0	19·794			
10·0	17·850			

angle at $R/S = 1{\cdot}64$ is 69·6° but is *less*, i.e. 69·4°, at $R/S = 1{\cdot}44$. Similar abnormalities, though of smaller size, occur in the regions of $R/S = 0{\cdot}85$, 0·60 and 0·45 ; in fact there are an infinite number of maxima and minima lying between here and $R = 0$.

Where a high degree of precision in measurement of mechanical wave velocity is desired, these peculiarities, the cause of which lies merely in the integration processes involved, must be taken into account. The predicted phase angle effects have been verified experimentally at these laboratories.

These phenomena, of course, are present in the case of any other radiation, such as electromagnetic, where two parallel coaxial apertures of appropriate size are in action.

ACKNOWLEDGMENTS

This work forms part of the Research Programme of the National Physical Laboratory and is published by permission of the Director. One of the authors (G. B.) desires to express his appreciation of the courtesy of Dr. Seki and Dr. Truell of Brown University in putting some field calculations at his disposal for a brief notice (Bradfield) of this effect given at the 3rd International Congress on Acoustics at Stuttgart in 1959.

REFERENCES

BRADFIELD, G., 1961, *Proc. 3rd Intern. Congr. Acoustics* (Amsterdam: Elsevier), p. 466.

HUNTINGDON, H. B., EMSLIE, A. G., and HUGHES, V. W., 1945, *J. Franklin Inst.*, **245**, 1.

SEKI, H., GRANATO, A., and TRUELL, R., 1956, *J. acoust. Soc. Amer.*, **28**, 230.

Wide Stacking Faults in High Purity Copper Crystals

By J. T. Fourie and R. J. Murphy
National Physical Research Laboratory, Pretoria, South Africa

[Received May 5, 1961]

During an investigation of plastic deformation in controlled orientation single-crystals of copper, which were grown from the melt by the technique of Chalmers (1953), we have obtained direct evidence for wide stacking faults.

Before being prepared for electron microscopy, the crystals had a cross section of 1 mm × 7 mm and were grown from high purity copper as supplied by Johnson, Matthey and Co. The total specified impurity content is 0·0008%, the main impurities being given as Ni, Fe, Ag, Ca, Mg and Si.

The orientations of the crystals were chosen so that after a crystal had been strained in tension, dislocations belonging to the primary system would be either in pure edge or in pure screw orientation when ending at the top and bottom surface of a thin foil. In a number of the samples of both crystal orientations, wide stacking faults were observed. Examples of these are shown in the plate. Wide stacking faults have only very rarely been observed in copper by direct transmission electron microscopy (Hirsch 1959). To our knowledge no suitable micrographs have appeared in the literature of stacking faults in copper crystals which had been prepared from bulk samples.

Upon observing the first set of wide stacking faults, the specific sample showing these faults was spectrographically analysed. It was found to be actually slightly purer than the original material, due to the zone refining effect of growing the crystal from the melt. The wide stacking faults could thus not be attributed to an impurity effect.

Stacking faults were never initially observed in the polished foil, but always formed during electron illumination. The stress required to split the partials was probably provided by the carbon contamination layer which builds up on the surface of the crystal. Similarly, the pile-up which can be seen in the plate also formed during the electron-microscope observation and not as a result of tensile deformation of the bulk specimen.

These observations show that under suitable stress conditions very wide stacking faults can be formed in thin copper crystals. Whether it is possible for these conditions of stress to exist in localized regions of bulk specimens is not certain.

References

Chalmers, B., 1953, *Canad. J. Phys.*, **31,** 136.
Hirsch, P. B., 1959, *J. Inst. Met.*, **87,** 408.

REVIEWS OF BOOKS

Solid State Physics, Vol. XI. Edited by F. SEITZ and D. TURNBULL. (New York and London: Academic Press, 1960.) [Pp. 438.] Price £4 9*s*. 6*d*.

THE contents: *Semiconducting Properties of Gray Tin* (2) by Busch and Kern; *Physics at High Pressures* (1) by Swenson; *The Effects of Elastic Deformation on the Electrical Conductivity of Semiconductors* (3) by Keyes; *Imperfection Ionization in CdS-Tyne Materials by Photoelectronic Techniques* (4) by Bube; *Cyclotron Resonance* (1) by Lax and Mavroides. The number after each title is its *coefficient of non-specificity*, calculated as follows: transform the title into a succession of A adjectives qualifying S substantives, omitting redundant words like ' physics ', ' effects ', ' properties ', ' materials ', etc. Then take the ratio A/S. Inspection of past and proposed titles shows quite clearly that if the coefficient is greater than 3 the article is too specialized for this general series. The optimum seems to be in the range $1 \leqslant A/S \leqslant 2$. This formula is freely offered to the editors, with the gracious request that they apply it rigorously in future.

J. M. Z.

The Hall Effect and Related Phenomena. By E. H. PUTLEY. (London: Butterworths, 1960.) [Pp. 263.] Price £2 10*s*. 0*d*.

" . . . . . . THIS book provides a review of the bulk transport properties of semiconductors and should be useful for research workers and students." It would be easy to dismiss it with some such bromide, and go on to something more interesting. Of its type it is not bad, so if I now appear harsh I hope Dr. Putley will take it that I am attacking a whole class, not his book in particular. It has the two major virtues of being written in good plain English, and of being clear and to the point. But what does he mean, in the preface, when he says " I have discussed the subject from the point of view of an experimentalist "? Not, surely, that it is all about potentiometers and cryostats, or that it is simply a set of tables of experimental data critically evaluated but without interpretation. No, this is a way of saying that he is *not to be taken too seriously when discussing the theory.*

Now, this modesty would be proper if the subject were full of inexplicable yet interesting phenomena, and the theories were obscure, unrealistic and ineffectual. A qualitative discussion of some of the models which had failed to work would then be just right. But semiconductor physics is well worked out in many aspects, and in such books as this we are always given long formulae for mobilities, magnetoresistance coefficients, thermoelectric effects, etc. with some of the formal algebra leading to them—that is, page after page of nothing but mathematical theory. The " point of view of the experimentalist " then becomes an excuse for failing to get the theory quite correct, for ignoring the assumptions and approximations that went into it and for giving only meaninglessly condensed descriptions of the explanations that have been given of the discrepancies. This is a field where theory and experiment have marched in unison; it is futile to believe that you can make any progress by taking a set of pointer readings and then thumbing through a book or two for a formula to fit.

Books like this are not given enough care in writing. If Dr. Putley had devoted to it the attention to detail, the scrupulosity, the critical mind which he would have given, as a matter of course, to the experiments and theory of, say, ten of his ordinary published papers (which would have covered about as many pages) then we should have had indeed, in the other stock phrase, " a valuable contribution to the subject ". As it is, he is not to be blamed—only our current attitude which puts " doing original research " above its consolidation and explanation.

J. M. Z.

Lectures on Fluid Mechanics. By SYDNEY GOLDSTEIN, with special lectures by J. M. Burgers. Series *Lectures in Applied Mathematics,* Vol. II. (Interscience Publishers, 1960.) [Pp. 309.] £2 10*s*. 0*d*.

FROM the preface ; " These lectures were prepared for the Seminar in Applied Mathematics arranged by the American Mathematical Society and the University of Colorado at Boulder, Colorado, June 23–July 19, 1957."

" The aim of the Seminar, as stated in the prospectus, was to give mature mathematicians an opportunity to become familiar with some of the major sectors in applied mathematics. There was a request that some aspect of Magnetohydrodynamics be included. Twelve lectures (plus two special lectures) were allotted to Fluid Mechanics. Some members of the audience were, in fact, already familiar with Fluid Mechanics : some were not. I did not believe myself capable of producing in twelve lectures a justified sense of familiarity with Fluid Mechanics in those who did not have it, so from the first I decided to compromise. Notes were reproduced and circulated, and most of them (though not all) were discussed afterwards in the lectures. Chapter 11, on the dynamics of inviscid gases, was not circulated before the lectures : it was added afterwards.

" Nevertheless, many important topics had to be omitted entirely or almost entirely. The choice in such cases is necessarily to some extent a personal one. Among the topics omitted I may mention turbulence, stability, surface and tidal waves, airfoil theory, and large parts of the dynamics of inviscid gases. . . ."

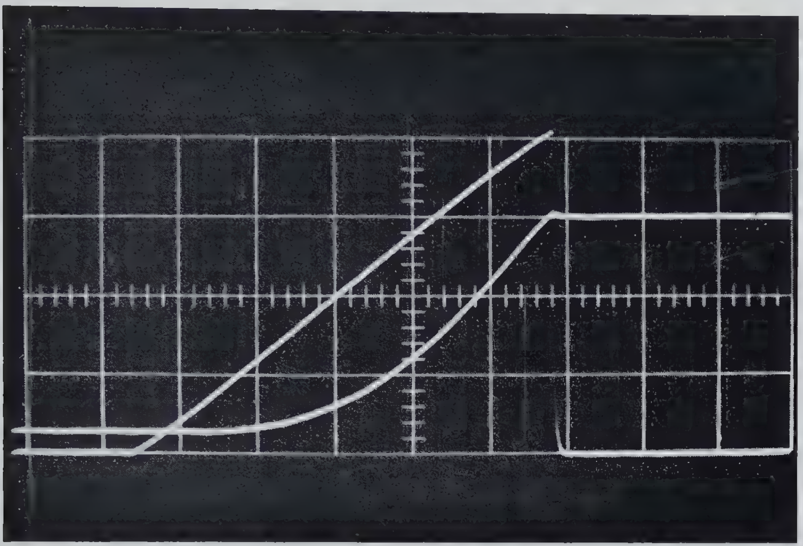
The chapter headings are 1. Kinematics, 2. Dynamics of the general fluid, 3. Electric and magnetic forces, 4. Inviscid fluids, 5. Newtonian fluids and the Navier–Stokes equations, 6. Exact solutions of the Navier–Stokes equations, 7. Boundary-layer theory for incompressible fluids, 8. Extensions of boundary-layer theory for incompressible fluids, 9. Boundary layers in gases, 10. Turbulence: Stability, 11. Dynamics of inviscid gases, 12. Mixtures and high temperature effects in gases, 13. Longitudinal (electrostatic) waves in an ionized gas. There are also four appendices. The last 40 pages contain the special lectures by J. M. Burgers entitled : Some problems of magneto-gas dynamics.

Whatever the ostensible purpose of these lectures, they will have a strong appeal to the specialist in fluid mechanics. There is room in this review only for some brief remarks on a few topics. The author has chosen (and this is noteworthy) to devote the first five chapters to a careful derivation of equations of motion. In chapters 7 and 8 he considers in some detail the singular perturbation problems of boundary-layer theory, and in particular the unsolved problems connected with the flow past a semi-infinite flat plate. For lack of space, turbulence and stability had to be dismissed with three pages of references. In chapter 11 the author considers subsonic and supersonic flow past slender bodies (with further notes in an appendix) and discusses the numerous unsolved problems connected with the validity of this approximation. The choice of topic and the precise style give implicitly an interesting picture of an eminent applied mathematician's view of applied mathematics.

F. U.

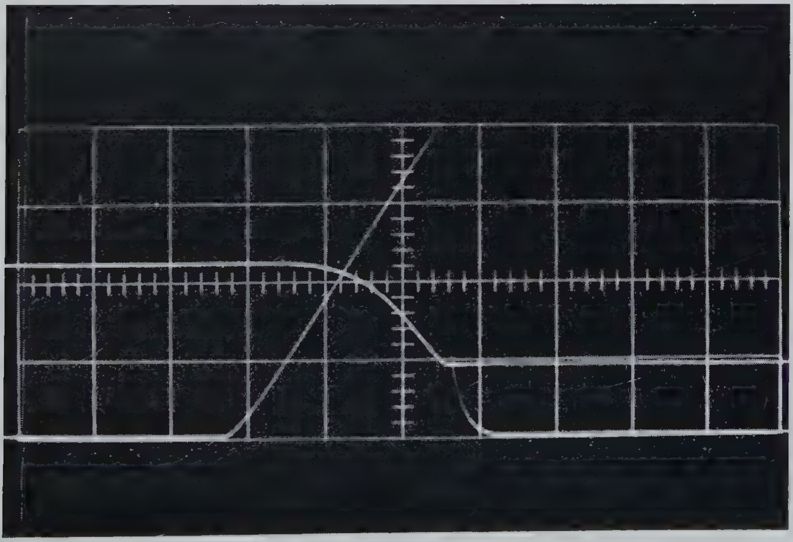
[*The Editors do not hold themselves responsible for the views expressed by their correspondents.*]

Fig. 1



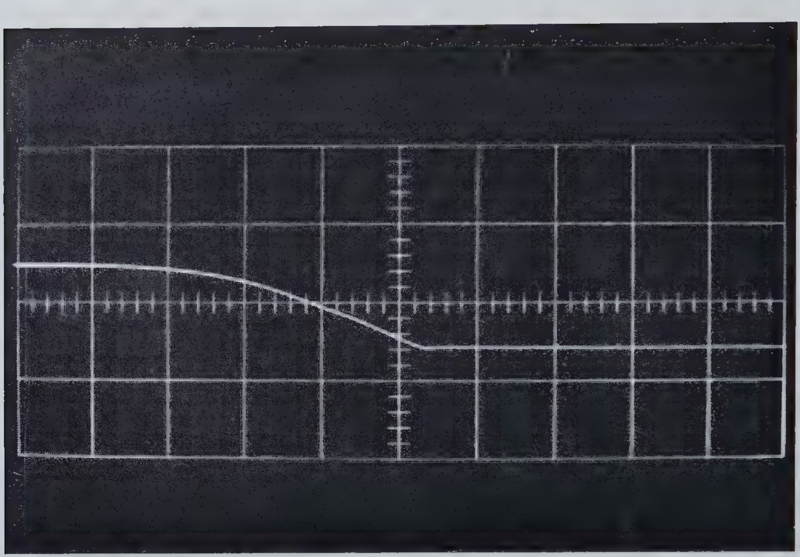
Oscilloscope trace of the saw-tooth current pulse and the obtained electrical effect. The duration of the pulse is 0·1 sec.

Fig. 2



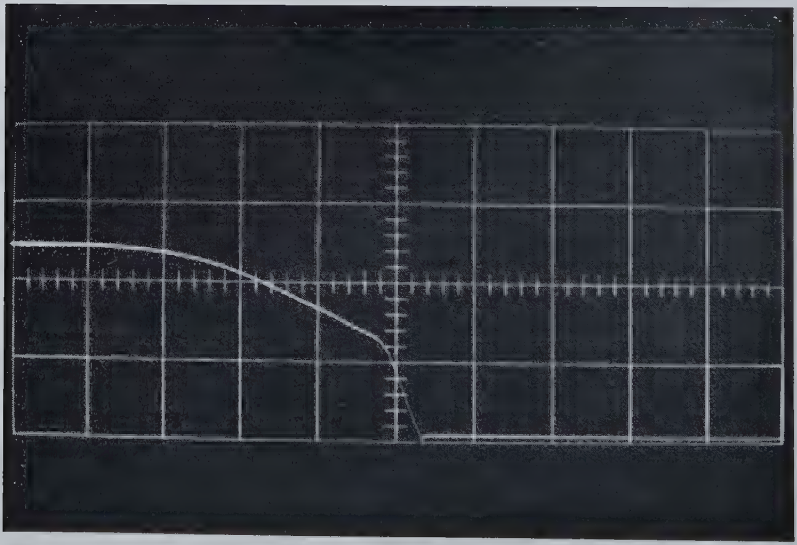
A saw-tooth current pulse with the corresponding signal. The duration of this pulse is ten times smaller than that of fig. 1.

Fig. 4



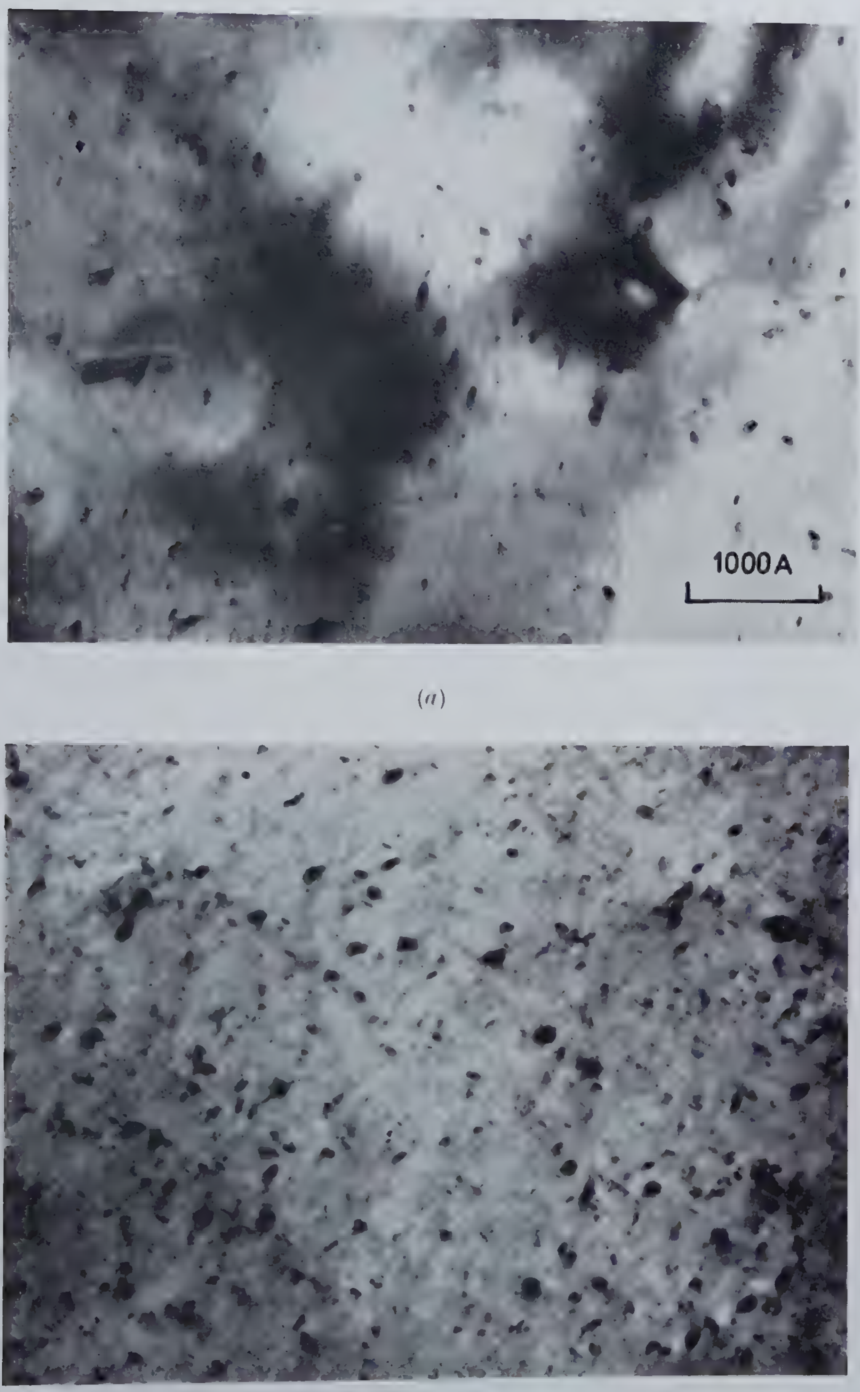
The electrical signal before the increase of the amplitude of deformation.

Fig. 5



The signal just after the increase of the amplitude of deformation.

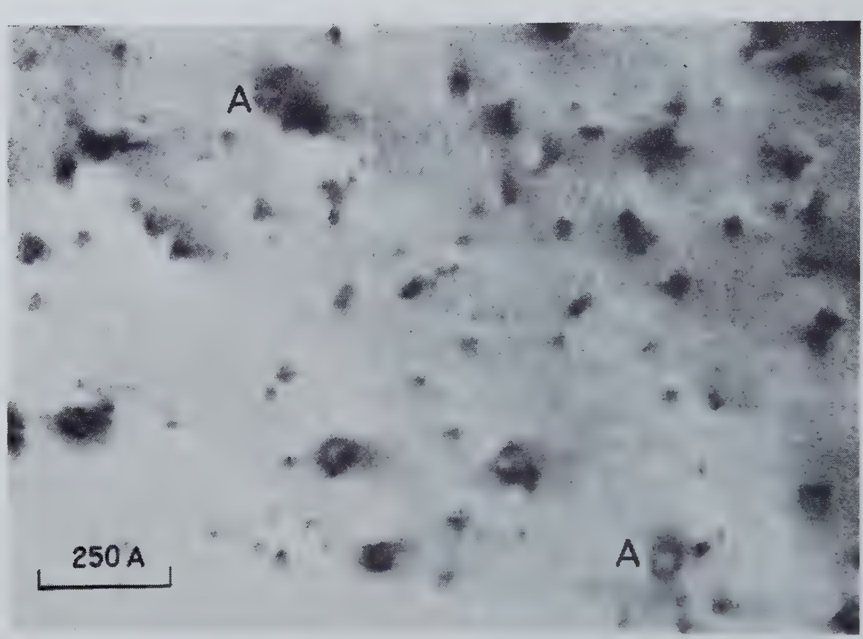
Fig. 1



(b)

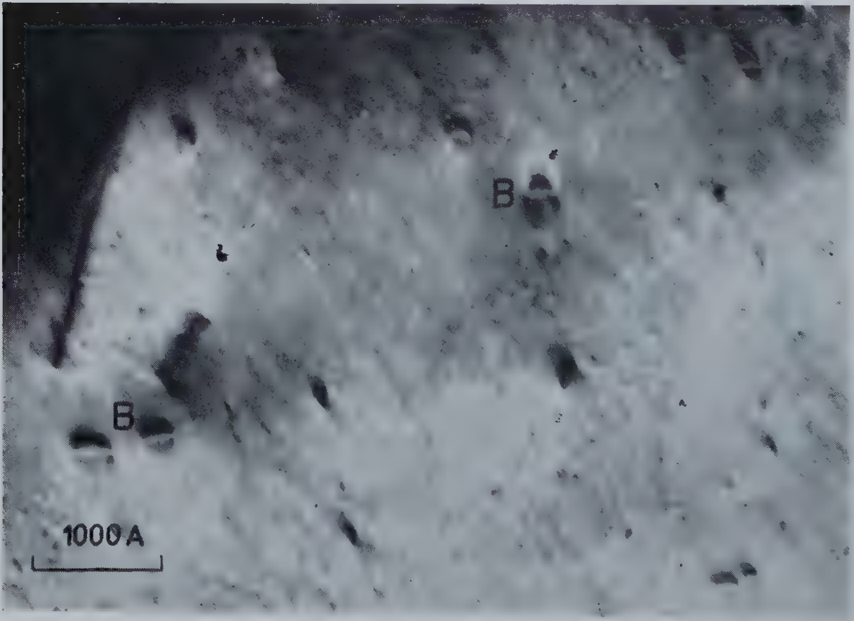
Au (111) after irradiation for (a) 100 sec, (b) 1000 sec. × 200 000.

Fig. 2



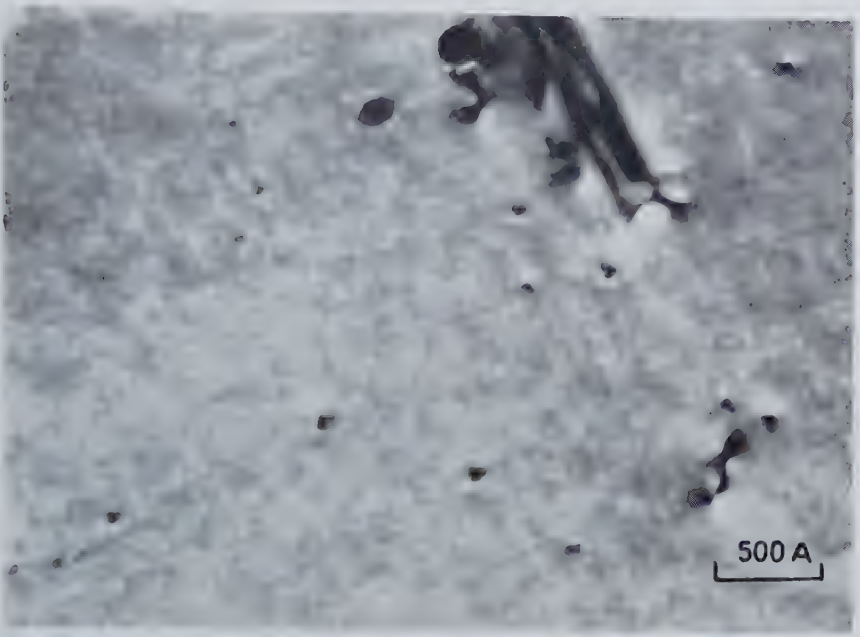
Characteristic appearance of the spots in a thin (100) Au film. ×800 000.

Fig. 3



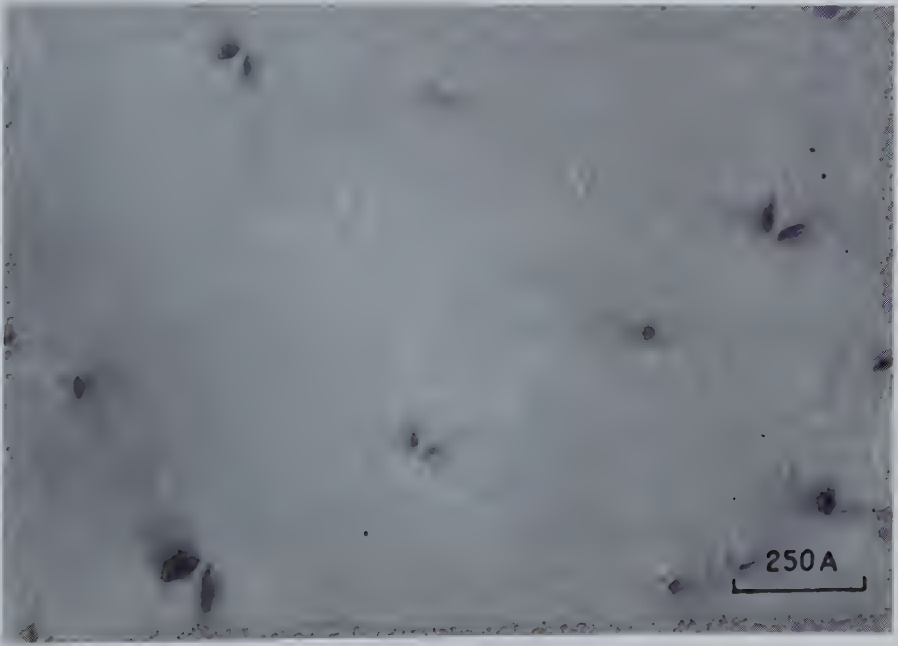
Large dislocation loops formed by irradiating (111) Au at ~300°c. ×200 000.

Fig. 4



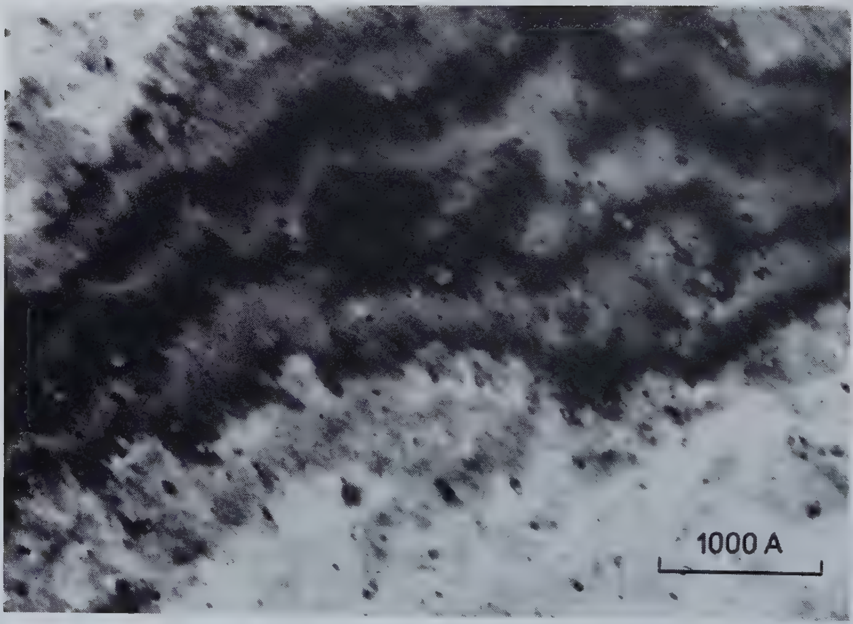
Micrographs showing the annealing out of spots at ~400°c in a (100) Au film (the remaining spot features are believed to be small tetrahedra of stacking fault). Dark field on (002). × 320 000.

Fig. 5



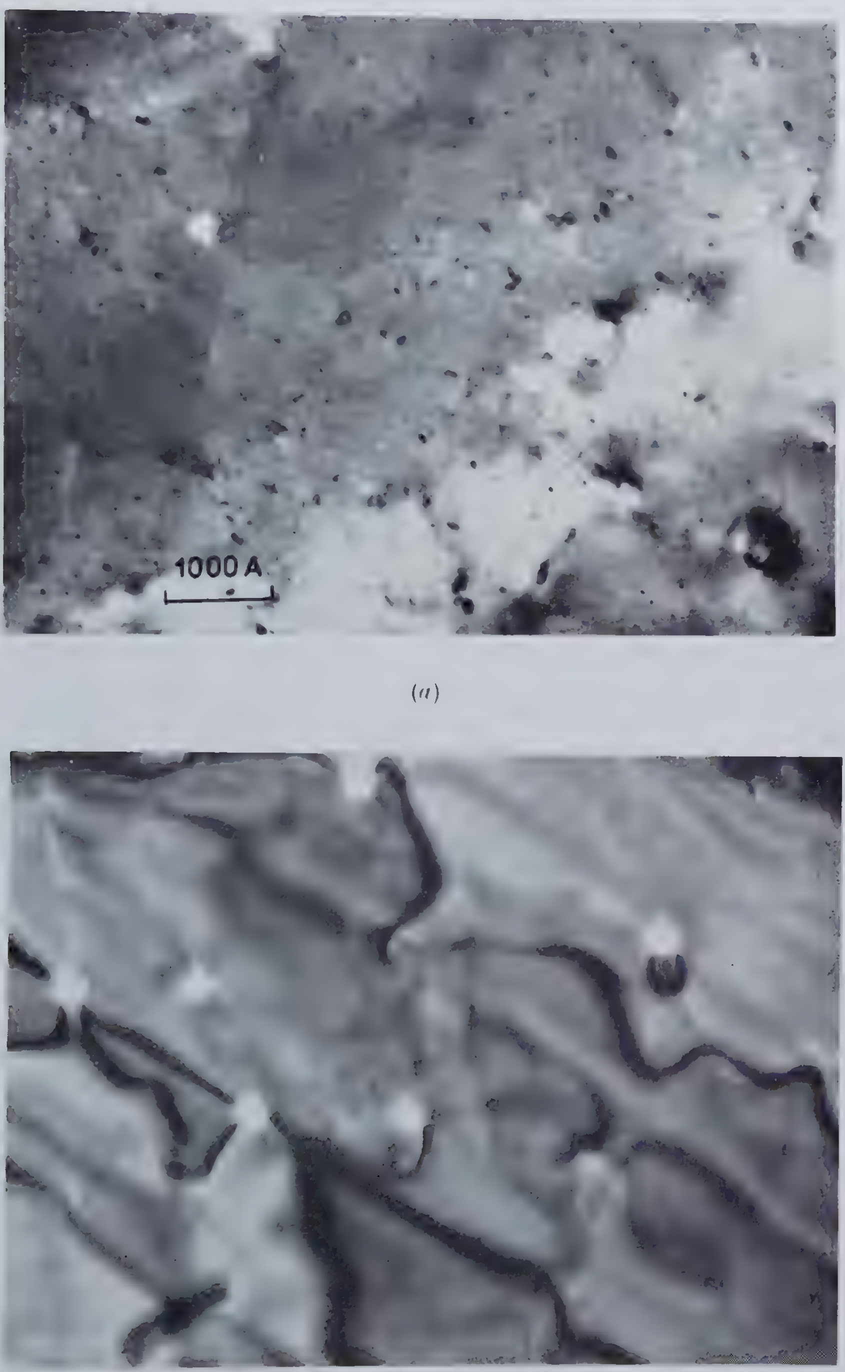
Tetrahedra of stacking fault produced in a (111) Au film irradiated at ~400°c. × 800 000.

Fig. 7



Loops and spots in a thin foil prepared by electropolishing bulk Au. × 240 000.

Fig. 8

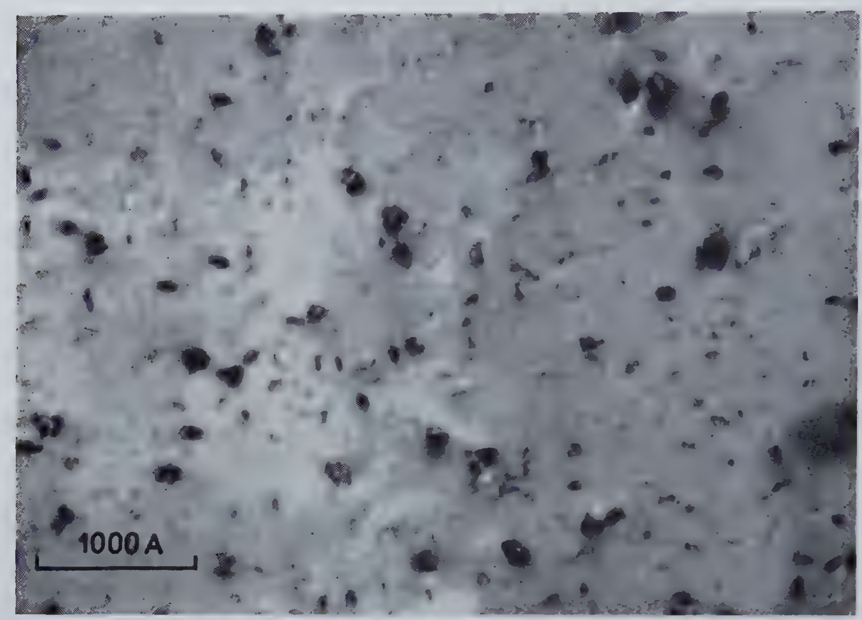


(a)

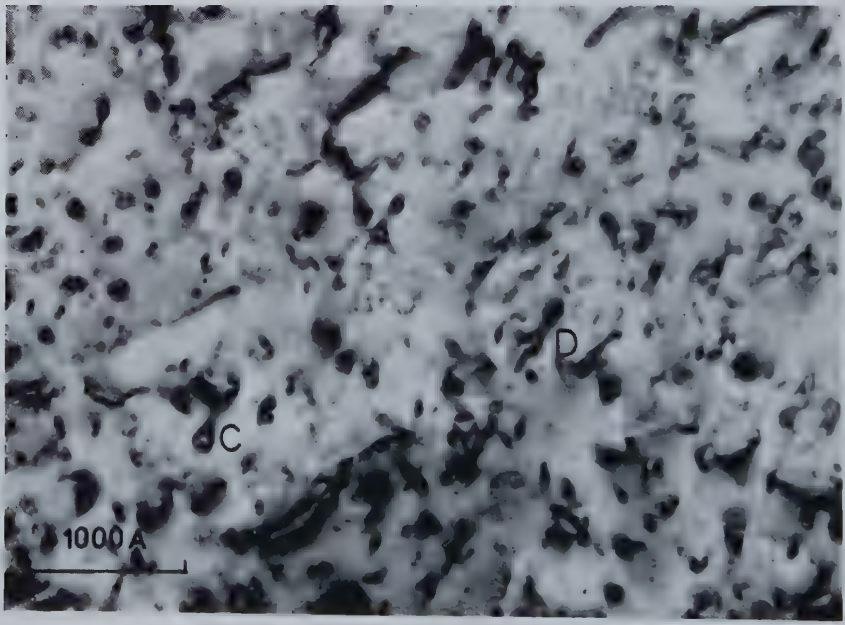
(b)

Micrographs obtained from two-film experiment, after about one hour's irradiation. (a) upper film; (b) lower film. ×160 000.

Fig. 9

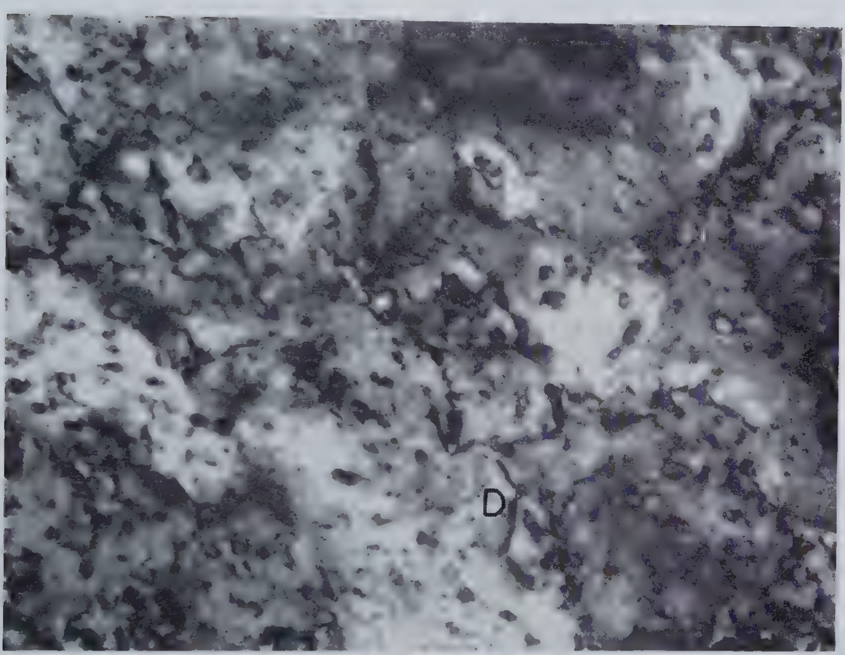


(*a*)



(*b*)

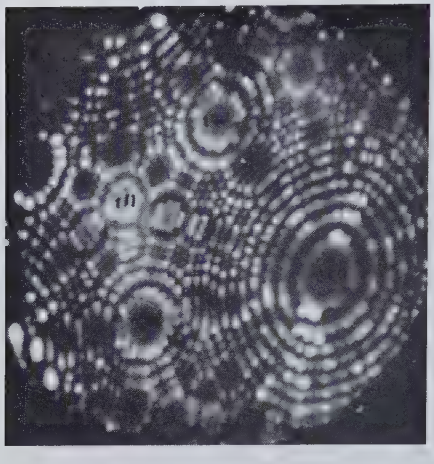
Fig. 9 (*continued*)



(*c*)

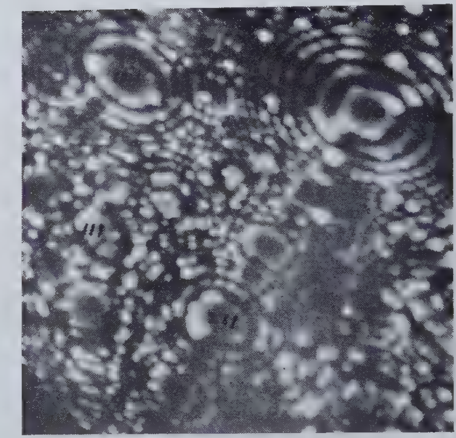
Damage produced in a (111) Au film, using oxide-coated filament. (*a*) light dose, (*b*) medium dose, (*c*) heavy dose. ×240 000.

Fig. 1



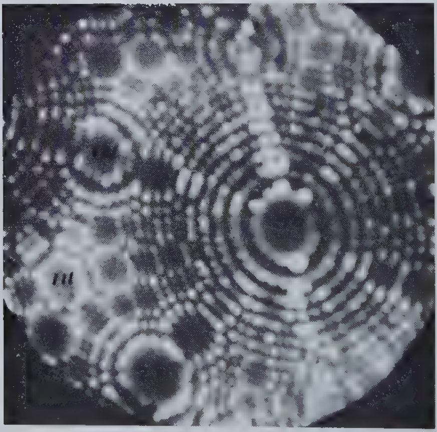
Tungsten. 9·5 kv, 5×10^{-3} mm. Helium. Resolution of atoms in the {211} planes is better in the ⟨110⟩ direction than in the ⟨111⟩ direction. Best resolution is obtained near the [111] pole.

Fig. 3



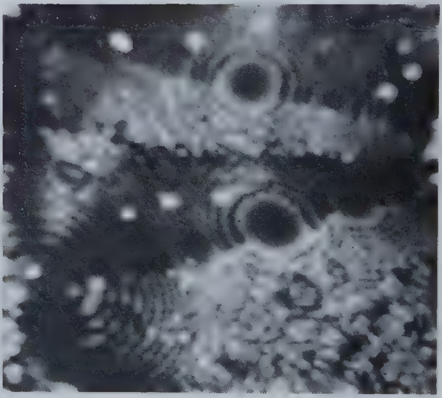
Molybdenum. 7 kv, 7×10^{-3} mm. Helium. Showing a bright triangular region around the [111] pole, with its apex at the (110) plane.

Fig. 4



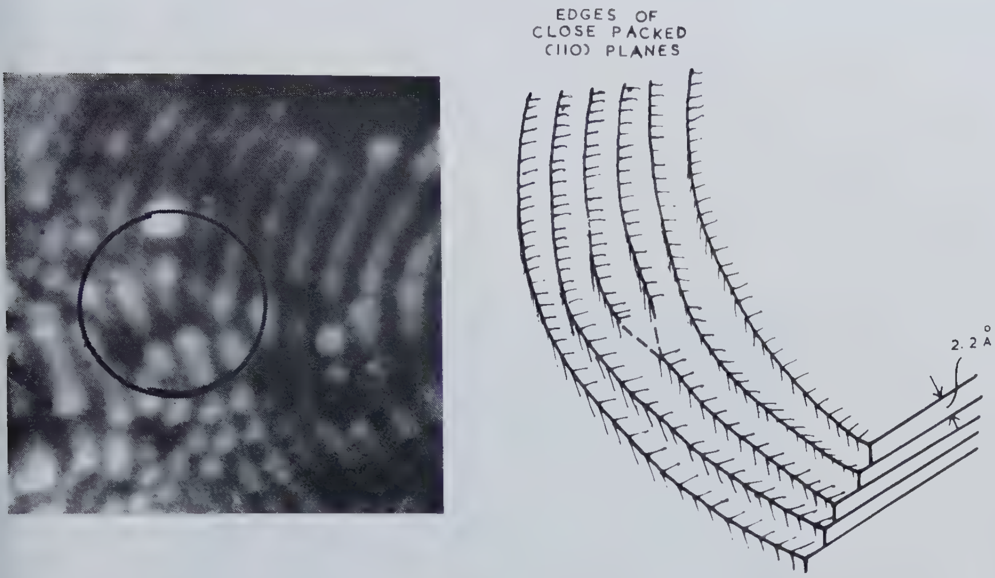
Tungsten. 10·8 kv, showing a bright band between ⟨111⟩ poles. The bright row of atoms running from [110] towards [010] is associated with a low coordination for the surface atoms in this region.

Fig. 5



Grain boundary in molybdenum. The large white spots are impurity atoms. The sharp changes in contrast are associated with the bright triangular region around the ⟨111⟩ poles in molybdenum (fig. 3).

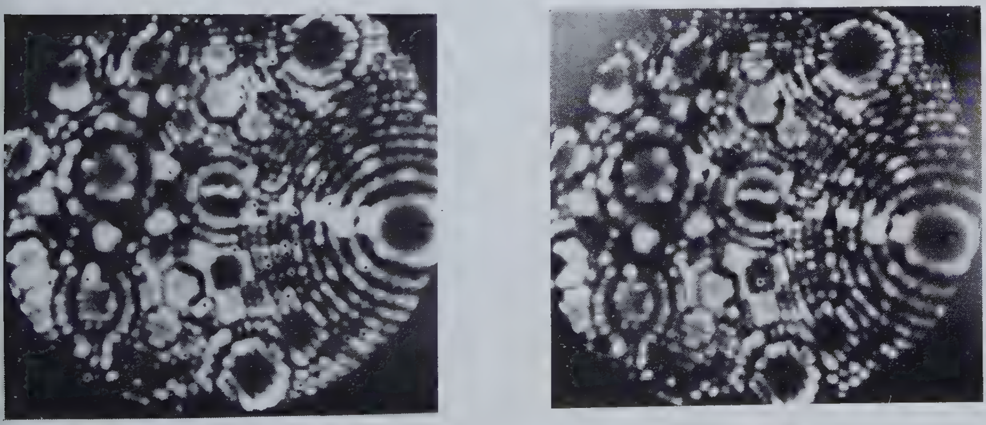
Fig. 8



(*a*) (*b*)

A dislocation in molybdenum in the region of the close-packed (110) planes. (*b*) Analysis of the dislocation in fig. 8(*a*) assuming the removal of half a (110) plane followed by a collapse and shear process.

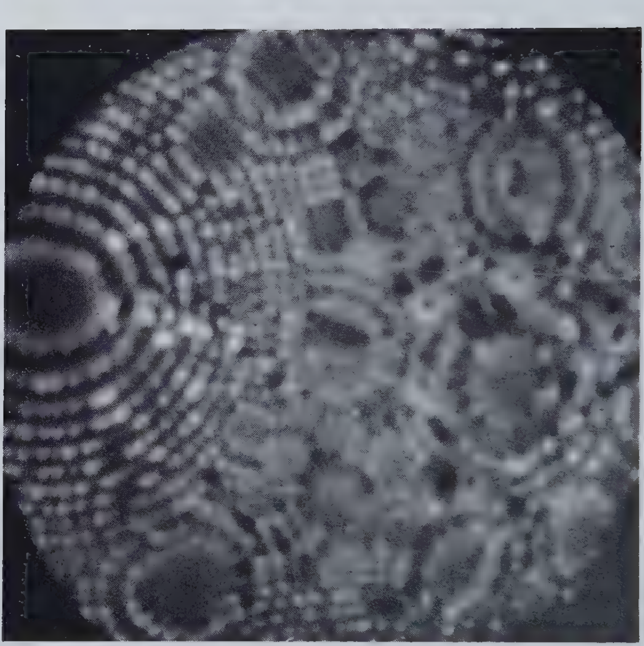
Fig. 10



(*a*) (*b*)

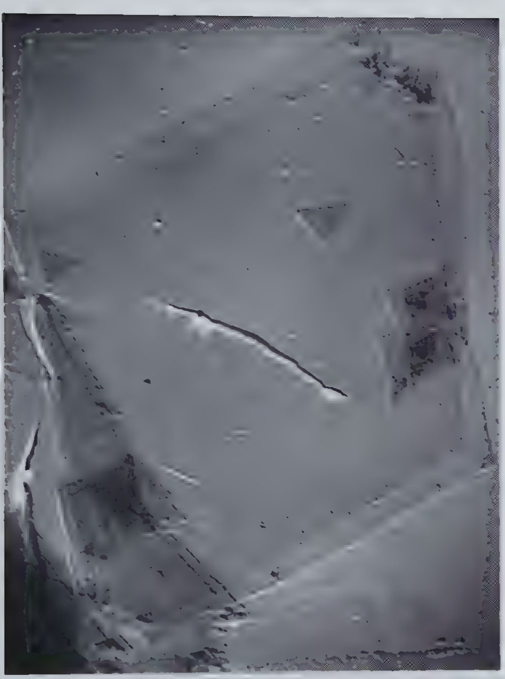
Tungsten irradiated with 5 MeV α-particles before (*a*) and after an impact (*b*). Atoms subsequently sputtered are dotted in fig. 10 (*a*) while two additional atoms are similarly marked in fig. 10 (*b*).

Fig. 11



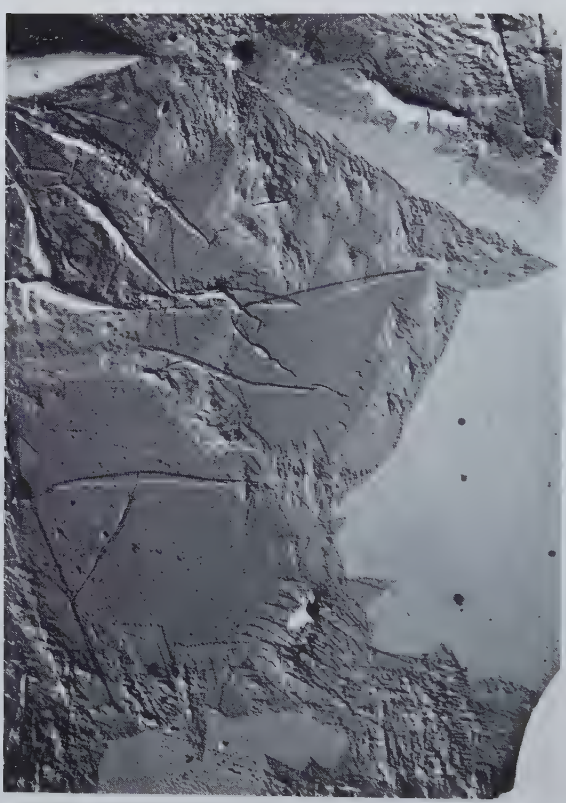
The effect of increasing the field to near the desorption region. Surface vacancies are clearly revealed.

Fig. 2



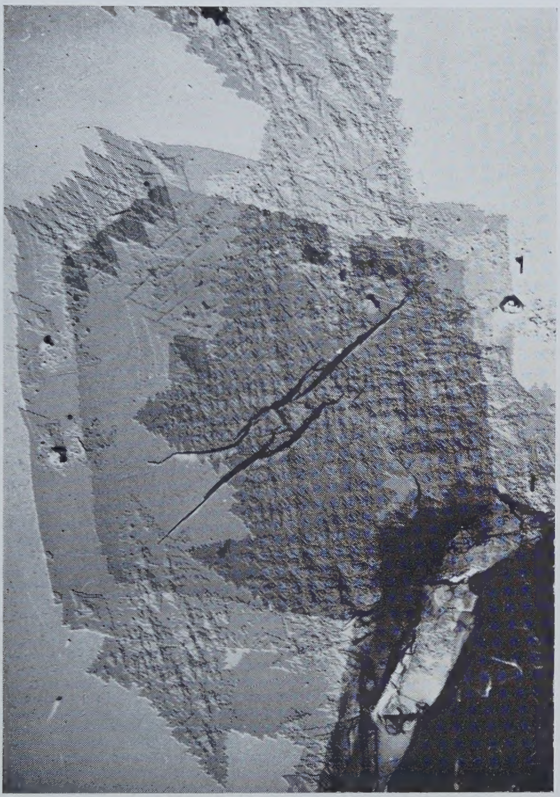
Conventional transmission electronmicrograph (80 kv) of a complex polyethylene crystal formed as described in the text. The original 150 Å thick truncated lozenge crystal has had grown on it a border only 110 Å thick, which has produced characteristic features in the crystal habit. (×4200.)

Fig. 3



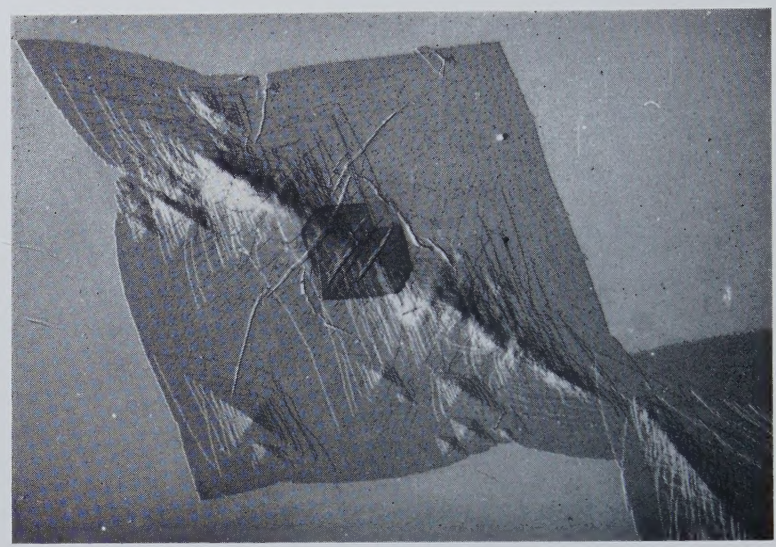
Detachment replica at 103°c in xylene of two polyethylene crystals as in fig. 2 in twin relation. Note that there are only two contrast tones over the crystals depending on whether a 150 or a 110 Å thick layer is uppermost. Electronmicrograph at 40 kv ; Au–Pd shadowed. (×2500.)

Fig. 4



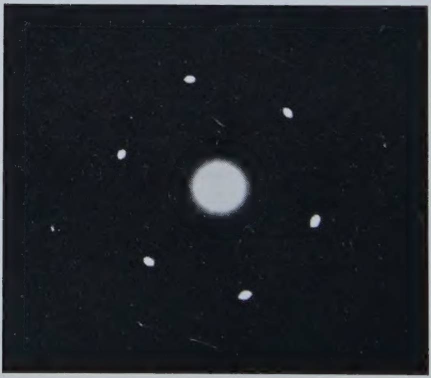
Detachment replica at 92°c in xylene of a polyethylene crystal as in fig. 2, overlain with and overlying dendritic crystals 110 Å thick. The contrast tones are due to the surface layer (cf. fig. 3) plus all layers 150 Å thick as these have been left undissolved. Electronmicrograph at 40 kv; Au–Pd shadowed. (× 65 000.)

Fig. 5

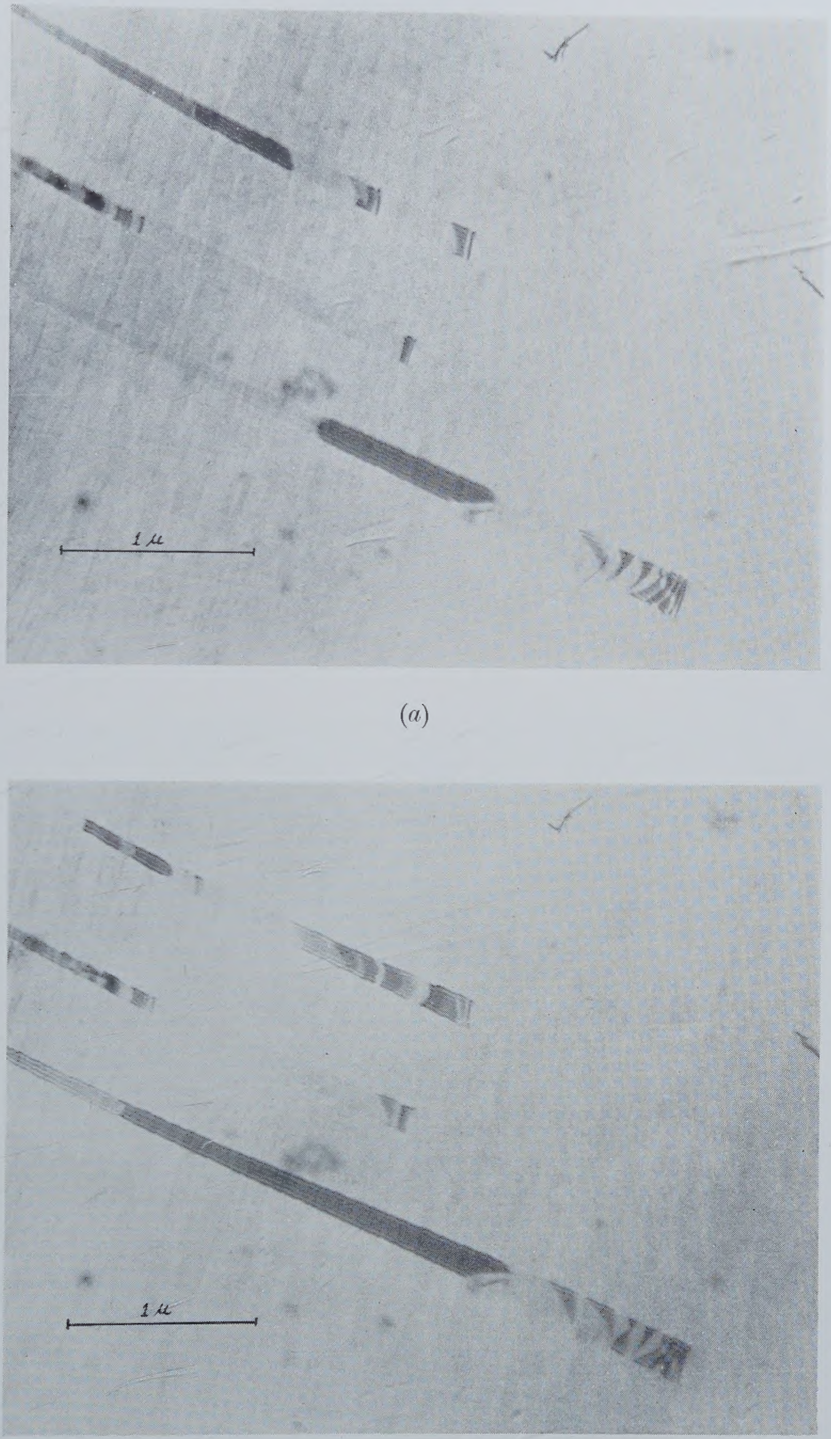


Detachment replica at 92°c in xylene of a multilayer polyethylene crystal, growing with 120 Å thick layers around a 150 Å thick monolayer truncated polyethylene lozenge seed. The contrast is of the type of fig. 4. Electronmicrograph at 40 kv ; Au–Pd shadowed. (×3200.)

Fig. 6



Electron diffraction pattern at 80 kv of a detachment replica similar to fig. 5 in which some of the layers were rotated slightly from the mean position.



(*a*)

(*b*)

(*a*) Wide stacking faults and a pile-up in the case of pure screw dislocations in a thin crystal of copper. (*b*) A further separation of some of the partials has occurred.